

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 10:13:23 : Search time 373 Seconds
(without alignments)
282.341 Million cell updates/sec

Title: US-09-904-603-1
Perfect score: 625
Sequence: 1 MSPDRPFKORRSFADRCKEY.....EKDEDFLYMVYASQETPGF 121

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
1:	/cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*		
2:	/cgn2_6/ptodata/1/paa/US06_COMB.pep.*		
3:	/cgn2_6/ptodata/1/paa/US07_COMB.pep.*		
4:	/cgn2_6/ptodata/1/paa/US08_COMB.pep.*		
5:	/cgn2_6/ptodata/1/paa/US081_COMB.pep.*		
6:	/cgn2_6/ptodata/1/paa/US082_COMB.pep.*		
7:	/cgn2_6/ptodata/1/paa/US083_COMB.pep.*		
8:	/cgn2_6/ptodata/1/paa/US084_COMB.pep.*		
9:	/cgn2_6/ptodata/1/paa/US085_COMB.pep.*		
10:	/cgn2_6/ptodata/1/paa/US087_COMB.pep.*		
11:	/cgn2_6/ptodata/1/paa/US088_COMB.pep.*		
12:	/cgn2_6/ptodata/1/paa/US089_COMB.pep.*		
13:	/cgn2_6/ptodata/1/paa/US090_COMB.pep.*		
14:	/cgn2_6/ptodata/1/paa/US091_COMB.pep.*		
15:	/cgn2_6/ptodata/1/paa/US092_COMB.pep.*		
16:	/cgn2_6/ptodata/1/paa/US093_COMB.pep.*		
17:	/cgn2_6/ptodata/1/paa/US094_COMB.pep.*		
18:	/cgn2_6/ptodata/1/paa/US095_COMB.pep.*		
19:	/cgn2_6/ptodata/1/paa/US096_COMB.pep.*		
20:	/cgn2_6/ptodata/1/paa/US097A_COMB.pep.*		
21:	/cgn2_6/ptodata/1/paa/US097B_COMB.pep.*		
22:	/cgn2_6/ptodata/1/paa/US098_COMB.pep.*		
23:	/cgn2_6/ptodata/1/paa/US099A_COMB.pep.*		
24:	/cgn2_6/ptodata/1/paa/US099B_COMB.pep.*		
25:	/cgn2_6/ptodata/1/paa/US100_COMB.pep.*		
26:	/cgn2_6/ptodata/1/paa/US101_COMB.pep.*		
27:	/cgn2_6/ptodata/1/paa/US102_COMB.pep.*		
28:	/cgn2_6/ptodata/1/paa/US103_COMB.pep.*		
29:	/cgn2_6/ptodata/1/paa/US104_COMB.pep.*		
30:	/cgn2_6/ptodata/1/paa/US105_COMB.pep.*		
31:	/cgn2_6/ptodata/1/paa/US106_COMB.pep.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match length DB ID	Description
1	625	100.0	121 24 US-09-904-603-1 Sequence 1, Appl 1
2	625	100.0	169 1 PCT-US01-18569-2536 Sequence 2536, Ap

3	625	100.0	169	28	US-10-264-049-2536	Sequence 2536, Ap
4	625	100.0	199	1	PCT-US01-08631-40152	Sequence 40152, A
5	621	99.4	181	30	US-10-405-027-4118	Sequence 4118, Ap
6	557	89.1	293	31	US-60-212-564-410	Sequence 410, App
7	545	87.2	107	23	US-09-834-366-14523	Sequence 14523, A
8	545	87.2	107	31	US-60-197-873-14523	Sequence 14523, A
9	521	83.4	123	23	US-09-834-366-15699	Sequence 15699, A
10	521	83.4	123	31	US-60-197-873-15699	Sequence 15699, A
11	521	83.4	125	18	US-09-488-725A-2805	Sequence 2805, Ap
12	521	83.4	125	23	US-09-834-366-20036	Sequence 20036, Ap
13	521	83.4	125	28	US-10-219-051B-7991	Sequence 7991, Ap
14	521	83.4	125	28	US-10-219-051B-13087	Sequence 13087, A
15	521	83.4	125	31	US-60-197-873-20036	Sequence 20036, A
16	521	83.4	142	13	US-08-944-467-7	Sequence 7, Appl 1
17	521	83.4	142	24	US-09-904-603-3	Sequence 3, Appl 1
18	521	83.4	142	28	US-10-205-194-71	Sequence 71, Appl 1
19	521	83.4	142	28	US-10-219-051B-13085	Sequence 13085, A
20	521	83.4	155	18	US-09-488-725A-6377	Sequence 6377, Ap
21	521	83.4	161	23	US-09-834-366-24980	Sequence 24980, A
22	521	83.4	161	31	US-60-197-873-24980	Sequence 24980, A
23	521	83.4	171	1	PCT-US00-05882-1051	Sequence 1051, Ap
24	521	83.4	171	24	US-09-925-301-1051	Sequence 1051, Ap
25	516	82.6	119	22	US-09-791-537-111211	Sequence 111211, A
26	516	82.6	141	28	US-10-219-051B-7989	Sequence 7989, Ap
27	466.5	74.6	99	1	PCT-US01-08631-40153	Sequence 40153, A
28	433	69.3	88	23	US-09-834-366-25917	Sequence 25917, A
29	433	69.3	88	31	US-60-197-873-25917	Sequence 25917, A
30	411	65.8	222	1	PCT-US01-08656-5682	Sequence 5682, Ap
31	400	64.0	98	23	US-09-621-976-7034	Sequence 7034, Ap
32	400	64.0	98	20	US-09-834-366-20032	Sequence 20032, A
33	400	64.0	98	31	US-60-147-499-7034	Sequence 7034, Ap
34	400	64.0	98	31	US-60-197-873-20032	Sequence 20032, A
35	381.5	61.0	130	22	US-09-791-537-6190	Sequence 6190, Ap
36	365	58.4	109	28	US-10-221-276-11949	Sequence 11949, Ap
37	359	57.4	147	22	US-09-786-797B-10	Sequence 10, Appl 1
38	359	57.4	147	31	US-60-131-321-8	Sequence 8, Appl 1
39	359	57.4	147	31	US-60-443-556-4508	Sequence 4508, Ap
40	359	57.4	147	31	US-60-455-444-8178	Sequence 8178, Ap
41	359	57.4	147	31	US-60-465-241-8178	Sequence 8178, Ap
42	349	55.8	167	22	US-09-758-472-7597	Sequence 7597, Ap
43	349	55.8	167	28	US-10-235-926-7597	Sequence 7597, Ap
44	317	50.7	82	22	US-09-758-450-564	Sequence 564, Ap
45	317	50.7	82	28	US-10-227-582-564	Sequence 564, Ap

ALIGNMENTS

RESULT 1
US-09-904-603-1
Sequence 1, Application US/09904603
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
G011, Surya K.
TITLE OF INVENTION: NOVEL MICROBUTYLE-ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Inocyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/904,603
FILING DATE: 12-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/905,117
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0211 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 121 amino acids
TYPE: amino acid
~STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: THVNOT03
CLONE: 144378

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

--
-- JS-09-904-603-1

	Query Match	100.0%	Score 625:	DB 24:	Length 121:
	Best Local Similarity	100.0%:	Pred. NO. 2.2e-65:		
	Matches 121:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0
QY	1	MPSDRPFKORSFPADCKEYQOIRDOHSHSKPIVIERRYKGKQLPVLDTKFLVPDHYNM	60		
Dd	1	MPSORPRFKQRSFPADCKEYQOIRDOHSHSKPIVIERRKGEKQLPYLDKTKFLVDHYNM	60		
QY	61	SELVKIIRRLQLNPTQAFFLLVNQHSWVSSTPIADIYEOKEDBGFLYYVASQETFG	120		
Dd	61	SELVKIIRRLQLNPTQAFFLLVNQHSWVSSTPIADIYEOKEDBGFLYYVASQETFG	120		
QY	121 F 121				
Dd	121 F 121				

```

RESULT 2
PCT-US01-18569-2536
: Sequence 2536 Application PC/TUS0118569
:
: GENERAL INFORMATION:
: APPLICANT: Human Genome Sciences, Inc.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA133POT
: CURRENT APPLICATION NUMBER: PCT/US01/18569
: CURRENT FILING DATE: 2001-06-07
: PRIOR APPLICATION NUMBER: 60/209,467
: PRIOR FILING DATE: 2000-06-07
: NUMBER OF SEQ ID NOS: 4360
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2536
:
: LENGTH: 169
: TYPE: PRT
: ORGANISM: Homo sapiens
: PCT-US01-18569-2536 -

```

Query Match	100.0%	Score 625	DB 1	Length 169
Best Local Similarity	100.0%	Pred. No. 3.6e-65		
Matches 121	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	MSDPRFKRSPFADCKEYVQOIRDOHSKIPVIERKYGKGEKOLPVLDTKFTFLVPDHYNM	60	
Db	49	MSDPRFKRSPFADCKEYVQOIRDOHSKIPVIERKYGKGEKOLPVLDTKFTFLVPDHYNM	106	
QY	61	SELVKIIRRRLOLNPQAFELLVNHNSMVSSTPIADIYEQEKDEGFLVNVYVASQETFG	120	
Db	109	SELVKIIRRRLOLNPQAFELLVNHNSMVSSTPIADIYEQEKDEGFLVNVYVASQETFG	168	
QY	121	F 121		
b	169	F 169		

```

RESULT 3
US-10-264-049-2536
: Sequence 2536, Application US/10264049
: GENERAL INFORMATION:
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PA1331
: CURRENT APPLICATION NUMBER: US/10/264,049
: PRIOR FILING DATE: 2002-10-04
: PRIOR APPLICATION NUMBER: PCT/US01/18569
: PRIOR FILING DATE: 2001-06-07
: PRIOR APPLICATION NUMBER: US 60/209,467
: PRIOR FILING DATE: 2000-06-07
: NUMBER OF SEQ ID NOS: 4360
: SOFTWARE: PatentIn Ver. 3.1
: SEQ ID NO 2536
: LENGTH: 169
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-264-049-2536

```

	Query Match	Similarity	Score	DB 28:	Length	169:
	Best Local	100.0%;	Pred. No.	3.6e-65;		
	Matches	Conservative	0;	Mismatches	0;	Gaps
	121;					
QY	1	MPSDRPFQKRRSPADRCKEVOQIRDOHSHSKIPVIERRYGGEKQLPVLDTKTFLVPDHYNM	60			
Dd	49	MPSDRPFGQRSPFADRCKEVOQIRQHRSKIPVIERRKGGEQLPVLDTKTFLVDHYNM	108			
QY	61	SELVKIIRRLQLNLTQAFFLLVNOSHWSVSTPIADIYEDEKEDGFLYYVASQETFG	120			
Dd	109	SELVKIIRRLQLNLTQAFFLLVNOSHWSVSTPIADIYEQEKEDGFLYYVASQETFG	168			
QY	121	F 121				
Dd	169	F 169				

```

RESULT 4
PCT-US01-08631-40152
Sequence 40152, Application PC/TUS0108631
GENERAL INFORMATION:
APPLICANT: Hyscq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-049
CURRENT APPLICATION NUMBER: PCT/US01/08631
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540, 217
PRIOR FILING DATE: 2000-09-31
PRIOR APPLICATION NUMBER: 09/649, 167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: custom
SEQ ID NO 40152
LENGTH: 199
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)...(199)
OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08631-40152

```

Query Match	100.0%	Score 625	DB 1	Length 199
Best Local Similarity	100.0%	Pred. NO. 4.6e-65		
Matches 121	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	MSDRPFKRRSFPADCKEVOQINDQHSKIPVILIERYKGGKQLPVLDKTFPLVPDHVM	60	
cb	52	MSDRPFKRRSFPADCKEVOQINDQHSKIPVILIERYKGGKQLPVLDKTFPLVPDHVM	111	


```
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 105
; OTHER INFORMATION: Xaa - Glu,Gln
US-60-197-873-14523

Query Match      87.2%; Score 545; DB 31; Length 107;
Best Local Similarity 99.1%; Pred. No. 5.8e-56;
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MPSDRPKRRSFADRCKEVQOIRDOHPKIPVIERKKGKQLPVLDKTKFLVPDHVM 60
    |||:|||||:|:|:|:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MPSEKTKQRRTFEQREVEDRLIREQHPKIPVIERKKGKQLPVLDKTKFLVPDHVM 60
    |||:|||||:|:|:|:|||||:|||||:|||||:|||||:|||||:|||||:

OY 61 SELVKIIRRLQLNPQAFLLVNGHSMVSSTPIADIYEQEKDEGFLVMVYASQETFG 120
    |||:|||||:|:|:|:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 SELKTIIRRLQLNANQAFLLVNGHSMVSSTPISEVSEKDEGFLVMVYASQETFG 120
    |||:|||||:|:|:|:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 9
US-09-834-366-15699
; Sequence 15699, Application US/09834366
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas, Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US2.REG
; CURRENT APPLICATION NUMBER: US/09/834,366
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 15699
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-366-15699

Query Match      83.4%; Score 521; DB 23; Length 123;
Best Local Similarity 82.5%; Pred. No. 5e-53;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

OY 1 MPSDRPKRRSFADRCKEVQOIRDOHPKIPVIERKKGKQLPVLDKTKFLVPDHVM 60
    |||:|||||:|:|:|:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MPSEKTKQRRTFEQREVEDRLIREQHPKIPVIERKKGKQLPVLDKTKFLVPDHVM 60
    |||:|||||:|:~|:|:|||||:|||||:|||||:|||||:|||||:|||||:

OY 61 SELVKIIRRLQLNPQAFLLVNGHSMVSSTPIADIYEQEKDEGFLVMVYASQETFG 120
    |||:|||||:|:|:|:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 SELKTIIRRLQLNANQAFLLVNGHSMVSSTPISEVSEKDEGFLVMVYASQETFG 120
    |||:|||||:|:~|:|:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 10
US-60-197-873-15699
; Sequence 15699, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas, Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US1.PRO
; CURRENT APPLICATION NUMBER: US/60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 15699
; LENGTH: 123
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-60-197-873-15699

Query Match      83.4%; Score 521; DB 31; Length 123;
Best Local Similarity 82.5%; Pred. No. 5e-53;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

OY 1 MPSDRPKRRSFADRCKEVQOIRDOHPKIPVIERKKGKQLPVLDKTKFLVPDHVM 60
    |||:|||||:|:~|:|:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MPSEKTKQRRTFEQREVEDRLIREQHPKIPVIERKKGKQLPVLDKTKFLVPDHVM 60
    |||:|||||:|:~|:|:|||||:|||||:|||||:|||||:|||||:|||||:

OY 61 SELVKIIRRLQLNPQAFLLVNGHSMVSSTPIADIYEQEKDEGFLVMVYASQETFG 120
    |||:|||||:|:~|:|:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 SELKTIIRRLQLNANQAFLLVNGHSMVSSTPISEVSEKDEGFLVMVYASQETFG 120
    |||:|||||:|:~|:|:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 11
US-09-488-725A-2805
; Sequence 2805, Application US/09488725A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784.FL.PCT
; CURRENT APPLICATION NUMBER: US/09/488,725A
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7144
; SOFTWARE: PL_FL_genes_b Versions 1.0
; SEQ ID NO 2805
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-488-725A-2805

Query Match      83.4%; Score 521; DB 18; Length 125;
Best Local Similarity 82.5%; Pred. No. 5.1e-53;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

OY 1 MPSDRPKRRSFADRCKEVQOIRDOHPKIPVIERKKGKQLPVLDKTKFLVPDHVM 60
    |||:|||||:|:~|:|:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MPSEKTKQRRTFEQREVEDRLIREQHPKIPVIERKKGKQLPVLDKTKFLVPDHVM 60
    |||:|||||:|:~|:|:|||||:|||||:|||||:|||||:|||||:|||||:

OY 61 SELVKIIRRLQLNPQAFLLVNGHSMVSSTPIADIYEQEKDEGFLVMVYASQETFG 120
    |||:|||||:|:~|:|:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 SELKTIIRRLQLNANQAFLLVNGHSMVSSTPISEVSEKDEGFLVMVYASQETFG 120
    |||:|||||:|:~|:|:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 12
US-09-834-366-20036
; Sequence 20036, Application US/09834366
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas, Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US2.REG
```

```

: CURRENT APPLICATION NUMBER: US/09/834,366
: CURRENT FILING DATE: 2001-04-13
: PRIOR APPLICATION NUMBER: US 60/197,873
: PRIOR FILING DATE: 2000-04-18
: NUMBER OF SEQ ID NOS: 52153
: SOFTWARE: Patent.pm
: SEQ ID NO: 20035
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-834-366-20036

Query Match
Best Local Similarity 83.4%; Score 521; DB 23; Length 125;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 MPSDRPFKORSFADRCKEVQIIRDOHPSKIPVITIERKKGKQLPVLKTKFLVDDHNM 60
DB 1 MPSEKTFKORRTFEGRVEDVRLIREQHPKIPVITIERKKGKQLPVLKTKFLVDDHNM 60
61 SELVKIIRRLQNLNFAFLVNOHSMVSVSTPIADIYEOEKDEDCGLYMWYASQETFG 120
|||||
61 SELIKIIRRLQNLNANQAFLLVNGHSMVSVSTPISEYSEKDEDCGLYMWYASQETFG 120

RESULT 13
US-10-219-051B-7991
: Sequence 7991, Application US/10219051B
: GENERAL INFORMATION:
: APPLICANT: The General Hospital Corporation doing business as Massachusetts General
: APPLICANT: Hospital / Bayer AG
: TITLE OF INVENTION: Nucleotide sequences involved in pain
: FILE REFERENCE: Lea 35693 Foreign Countries
: CURRENT APPLICATION NUMBER: US/10/219,051B
: CURRENT FILING DATE: 2003-05-09
: PRIOR APPLICATION NUMBER: US 60/312,147
: PRIOR FILING DATE: 2001-08-14
: PRIOR APPLICATION NUMBER: US 60/346,382
: PRIOR FILING DATE: 2001-11-01
: PRIOR APPLICATION NUMBER: US 60/333,347
: PRIOR FILING DATE: 2001-11-26
: NUMBER OF SEQ ID NOS: 14715
: SOFTWARE: Perl script
: SEQ ID NO: 7991
: LENGTH: 125
: TYPE: PRT
: ORGANISM: Homo sapiens
: PUBLICATION INFORMATION:
: DATABASE ACCESSION NUMBER: Refseq / NP_073729
: DATABASE ENTRY DATE: 2002-11-05
US-10-219-051B-7991

Query Match
Best Local Similarity 83.4%; Score 521; DB 28; Length 125;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 MPSDRPFKORSFADRCKEVQIIRDOHPSKIPVITIERKKGKQLPVLKTKFLVDDHNM 60
DB 1 MPSEKTFKORRTFEGRVEDVRLIREQHPKIPVITIERKKGKQLPVLKTKFLVDDHNM 60
61 SELVKIIRRLQNLNFAFLVNOHSMVSVSTPIADIYEOEKDEDCGLYMWYASQETFG 120
|||||
61 SELIKIIRRLQNLNANQAFLLVNGHSMVSVSTPISEYSEKDEDCGLYMWYASQETFG 120

RESULT 14
US-10-219-051B-13087
: Sequence 13087, Application US/10219051B
: GENERAL INFORMATION:
: APPLICANT: The General Hospital Corporation doing business as Massachusetts General
: APPLICANT: Hospital / Bayer AG
: TITLE OF INVENTION: Nucleotide sequences involved in pain
: FILE REFERENCE: Lea 35693 Foreign Countries

```

```

: CURRENT APPLICATION NUMBER: US/10/219,051B
: CURRENT FILING DATE: 2003-05-09
: PRIOR APPLICATION NUMBER: US 60/312,147
: PRIOR FILING DATE: 2001-08-14
: PRIOR APPLICATION NUMBER: US 60/346,382
: PRIOR FILING DATE: 2001-11-01
: PRIOR APPLICATION NUMBER: US 60/333,347
: PRIOR FILING DATE: 2001-11-26
: NUMBER OF SEQ ID NOS: 14715
: SOFTWARE: Perl script
: SEQ ID NO: 13087
: LENGTH: 125
: TYPE: PRT
: ORGANISM: Homo sapiens
: PUBLICATION INFORMATION:
: DATABASE ACCESSION NUMBER: Refseq / NP_073729
: DATABASE ENTRY DATE: 2002-11-05
US-10-219-051B-13087

Query Match
Best Local Similarity 83.4%; Score 521; DB 28; Length 125;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 MPSDRPFKORSFADRCKEVQIIRDOHPSKIPVITIERKKGKQLPVLKTKFLVDDHNM 60
DB 1 MPSEKTFKORRTFEGRVEDVRLIREQHPKIPVITIERKKGKQLPVLKTKFLVDDHNM 60
61 SELVKIIRRLQNLNFAFLVNOHSMVSVSTPIADIYEOEKDEDCGLYMWYASQETFG 120
|||||
61 SELIKIIRRLQNLNANQAFLLVNGHSMVSVSTPISEYSEKDEDCGLYMWYASQETFG 120

RESULT 15
US-60-197-873-20036
: Sequence 20036, Application US/60197873
: GENERAL INFORMATION:
: APPLICANT: Bejantin, Stephanie
: APPLICANT: Tanaka, Hiroaki
: APPLICANT: Dumas Milne Edwards, Jean Baptiste
: APPLICANT: Joebert, Severin
: APPLICANT: Giordano, Jean-Yves
: TITLE OF INVENTION: ESTs and Encoded Human Proteins.
: FILE REFERENCE: 81,051,PRO
: CURRENT APPLICATION NUMBER: US/60/197,873
: CURRENT FILING DATE: 2000-04-18
: NUMBER OF SEQ ID NOS: 52153
: SOFTWARE: Patent.pm
: SEQ ID NO: 20036
: LENGTH: 125
: TYPE: PRT
: ORGANISM: Homo sapiens
US-60-197-873-20036

Query Match
Best Local Similarity 83.4%; Score 521; DB 31; Length 125;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 MPSDRPFKORSFADRCKEVQIIRDOHPSKIPVITIERKKGKQLPVLKTKFLVDDHNM 60
DB 1 MPSEKTFKORRTFEGRVEDVRLIREQHPKIPVITIERKKGKQLPVLKTKFLVDDHNM 60
61 SELVKIIRRLQNLNFAFLVNOHSMVSVSTPIADIYEOEKDEDCGLYMWYASQETFG 120
|||||
61 SELIKIIRRLQNLNANQAFLLVNGHSMVSVSTPISEYSEKDEDCGLYMWYASQETFG 120

Search completed: August 11, 2003, 10:23:10
Job time : 375 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

```
Run on:      August 11, 2003, 10:13:53 ; Search time 21 Seconds
            (without alignments)
            115.496 Million cell updates/sec
```

Title: US-09-904-603-1
 Perfect score: 625
 Sequence: 1 MPSDRPFKQRRSFADRCREV.....EKEDGFLYMYASQETFGF 121

Scoring table: RI.OSUM62
Gapop 10.0 , Gapext 0.5

Searched: 98705 seqs, 20044758 residues

Total number of hits satisfying chosen parameters: 98705

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%
	Listing first 45 summaries

```
Database : Pending_Patents_AA_New:*
1 : /cgn2-6/ptodata/1/paa/PC1_NEW_COMB pep :*
2 : /cgn2-6/ptodata/1/paa/S06_NEW_COMB pep :*
3 : /cgn2-6/ptodata/1/paa/S07_NEW_COMB pep :*
4 : /cgn2-6/ptodata/1/paa/S08_NEW_COMB pep :*
5 : /cgn2-6/ptodata/1/paa/S09_NEW_COMB pep :*
6 : /cgn2-6/ptodata/1/paa/S10_NEW_COMB pep :*
7 : /cgn2-6/ptodata/1/paa/S16_NEW_COMB pep :*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	DB	ID	Description
1	521	83.4	125	6	US-10-286-897-2805	Sequence 2805, App
2	521	83.4	155	6	US-10-286-897-6377	Sequence 6377, App
3	411	65.8	222	6	US-10-273-573-6982	Sequence 6982, App
4	401	48.2	120	6	US-10-293-244-3825	Sequence 3825, App
5	272	43.5	135	6	US-10-293-244-1857	Sequence 1857, App
6	256	41.0	50	5	US-09-820-649-221	Sequence 221, App
7	235	37.8	46	5	US-09-820-649-222	Sequence 222, App
8	203	32.5	117	6	US-10-408-765a-1031	Sequence 1021, App
9	193	30.9	133	6	US-10-603-113-1457a	Sequence 1457a, App
10	68.5	11.0	602	6	US-10-451-901-4	Sequence 4, App
11	68.5	11.0	1648	6	US-10-603-113-14286	Sequence 14286, App
12	68	10.9	243	6	US-10-617-320-4624	Sequence 4624, App
13	68	10.9	1616	5	US-09-820-843b-16	Sequence 16, App
14	66.5	10.6	469	6	US-10-603-113-15278	Sequence 15278, App
15	66	10.6	454	6	US-10-603-113-17078	Sequence 17078, App
16	65.5	10.5	283	6	US-10-603-113-21629	Sequence 21629, App
17	65.5	10.5	322	6	US-10-326-956-267	Sequence 267, App
18	65.5	10.5	1178	6	US-10-326-956-2805	Sequence 2805, App
19	65.5	10.5	2000	6	US-10-326-956-339	Sequence 339, App
20	65.5	10.5	4910	6	US-10-346-863-21	Sequence 21, App
21	65	10.4	1047	6	US-10-346-956-2653	Sequence 2653, App
22	65	10.4	3058	7	US-60-487-610-2541	Sequence 2541, App
23	64.5	10.3	655	6	US-10-326-956-483	Sequence 483, App
24	64	10.2	308	6	US-10-294-433-233	Sequence 233, App
25	64.5	10.2	303	6	US-10-617-320-4703	Sequence 4703, App
26	63	10.1	45	5	US-09-820-649-220	Sequence 220, App

27	63	10.1	468	6	05-10-10-7-320-2827	Sequence 4672, A
28	63	10.1	472	6	05-10-60-8-113-19832	Sequence 17572, A
29	63	10.1	588	6	05-10-408-765A-907	Sequence 707, App
30	62.5	10.0	283	6	05-10-408-417A-50	Sequence 707, App
31	62.5	10.0	283	6	05-10-7-27-8-578-5678	Sequence 5778, A
32	62.5	10.0	327	6	05-10-61-7-320-333	Sequence 3331, A
33	62.5	10.0	380	6	05-10-27-8-578-5513	Sequence 5513, App
34	62.5	10.0	444	6	05-10-60-8-114-4924	Sequence 4924, App
35	62.5	10.0	539	6	05-10-294-433-393	Sequence 1933, App
36	62	9.9	309	6	05-10-526-955-1507	Sequence 1507, App
37	62	9.9	527	6	05-10-60-8-114-4374	Sequence 4374, App
38	62	9.9	1098	6	05-10-326-956-393	Sequence 393, App
39	62	9.9	1098	6	05-10-61-3-520-2153	Sequence 2153, A
40	61.5	9.8	307	6	05-10-60-8-113-20251	Sequence 20251, A
41	61.5	9.8	475	6	05-10-326-956-1259	Sequence 1259, App
42	61.5	9.8	1151	7	05-60-478-194-3270	Sequence 4270, App
43	61.5	9.8	1311	6	05-10-408-765A-2541	Sequence 2541, App
44	61.5	9.8	2033	6	05-10-408-765A-419	Sequence 419, App
45	61	9.8	395	6	05-10-60-8-113-17865	Sequence 17865, A

ALIGNMENT

```

RESULT 1
US-10-286-897-2805
? Sequence 2805, Application US/10286897
? GENERAL INFORMATION:
? APPLICANT: Hyseq Inc
? TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
? FILE REFERENCE: 784FLPT
? CURRENT APPLICATION NUMBER: US/10/286,897
? CURRENT FILING DATE: 2002-11-01
? PRIOR APPLICATION NUMBER: US/09/4488,725
? PRIOR FILING DATE: 2000-01-21
? PRIOR APPLICATION NUMBER: US/09/552,317
? PRIOR FILING DATE: 2000-04-25
? PRIOR APPLICATION NUMBER: US/09/598,042
? PRIOR FILING DATE: 2000-06-20
? PRIOR APPLICATION NUMBER: US/09/620,312
? PRIOR FILING DATE: 2000-07-19
? PRIOR APPLICATION NUMBER: US/09/653,450
? PRIOR FILING DATE: 2000-08-31
? PRIOR APPLICATION NUMBER: US/09/662,191
? PRIOR FILING DATE: 2000-09-14
? PRIOR APPLICATION NUMBER: US/09/693,036
? PRIOR FILING DATE: 2000-10-19
? PRIOR APPLICATION NUMBER: US/09/727,344
? PRIOR FILING DATE: 2000-11-29
? NUMBER OF SEQ ID NOS: 7143
? SOFTWARE: PL_FL_genes_b Versions 1.0
? SEQ ID NO 2805
? LENGTH: 125
? TYPE: PRT
? ORGANISM: Homo sapiens
US-10-286-897-2805

Query Match      83.4%; Score 521; LH 6; Length 125;
Best Local Similarity   82.5%; Pred. No. 2,1e-51;
Matches    99; Conservative    12; Mismatches    9; Indels    0; Gaps    0;

QY          1  MPSRPRKGRSFSADKRGKYVQLIKIQHPSKIPVIIEBYKKRPKQIPLVDIKTKIFLVHLVNM    60
|||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
DB          1  MPSRKTRKQRKTFFQRFQDVRLREKHHTKIPIVIERKCHKQIPLVDIKTKIFLVGHVNM    60
|||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

QY          61  SELVKTIIRRLQLNPITAFLLVMQHSMVSSTPDIADIVQEKENDEFEELMYVASLSETF    120
|||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
DB          61  SELKIIRRLQLNANGAFLVMNHSMVSSTPISFVESKEKDGEFEELMYVASLSETF    120
|||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

RESULT 2
US-10-286-897-6377
? Sequence 6377, Application US/10286897

```

RESULT 2
US-10-286-897-6377
; Sequence 6377, Application US/10286897

```
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq Inc
;; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;; FILE REFERENCE: 784FLPCT
;; CURRENT APPLICATION NUMBER: US/10/286,897
;; CURRENT FILING DATE: 2002-11-01
;; PRIOR APPLICATION NUMBER: US/09/488,725
;; PRIOR FILING DATE: 2000-01-21
;; PRIOR APPLICATION NUMBER: US/09/552,317
;; PRIOR FILING DATE: 2000-04-25
;; PRIOR APPLICATION NUMBER: US/09/598,042
;; PRIOR FILING DATE: 2000-06-20
;; PRIOR APPLICATION NUMBER: US/09/620,312
;; PRIOR FILING DATE: 2000-07-19
;; PRIOR APPLICATION NUMBER: US/09/653,450
;; PRIOR FILING DATE: 2000-08-31
;; PRIOR APPLICATION NUMBER: US/09/662,191
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: US/09/693,036
;; PRIOR FILING DATE: 2000-10-19
;; PRIOR APPLICATION NUMBER: US/09/727,344
;; PRIOR FILING DATE: 2000-11-29
;; NUMBER OF SEQ ID NOS: 7143
;; SOFTWARE: pf-fl-genes_b Versions 1.0
;; SEQ ID NO: 6377
;; LENGTH: 155
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-286-897-6377

Query Match      83.4%; Score 521; DB 6; Length 155;
Best Local Similarity 82.5%; Pred. No. 2.8e-51;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 MSPDRPFKORSFADCKEVOQIRDOHPSKIPVIERYKGEKQLPVLDTKFLVDPDHYNM 60
    |||:|||||:| |::|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 31 MPSEKTFKQRTFEQRYVDVRLIREQHTKIPVIERKKGKQLPVLDTKFLVDPDHYNM 90
QY 61 SELVKIIRRLQLNPTQAFLLVNOHSMVSVSTPIADIYEOEKDEDEGFLYMYVASOETFG 120
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 91 SELIKIIRRLQLNANOAFLLVNGHSMVSVSTPISEYSEKEDDEGFLYMYVASOETFG 150

RESULT 3
US-10-273-573-6982
;; Sequence 6982, Application US/10273573
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq, Inc
;; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
;; FILE REFERENCE: 21272-066
;; CURRENT APPLICATION NUMBER: US/10/273,573
;; CURRENT FILING DATE: 2002-10-18
;; PRIOR APPLICATION NUMBER: 09/522,929
;; PRIOR FILING DATE: 2000-04-18
;; PRIOR APPLICATION NUMBER: 09/770,160
;; PRIOR FILING DATE: 2001-01-26
;; NUMBER OF SEQ ID NOS: 10994
;; SOFTWARE: Custom
;; SEQ ID NO: 6982
;; LENGTH: 222
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: DOMAIN
;; LOCATION: (9)...(31)
;; OTHER INFORMATION: PROSTANOID EPI RECEPTOR SIGNATURE domain identified by
;; OTHER INFORMATION: EMATRIX, accession number PR00580E, p-value=2.014e-10, raw score
;; OTHER INFORMATION: 7.15
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(222)
;; OTHER INFORMATION: Xaa = X or * as defined in Table 2.
US-10-273-573-6982
```

```
Query Match      65.8%; Score 411; DB 6; Length 222;
Best Local Similarity 67.2%; Pred. No. 9.1e-39;
Matches 82; Conservative 16; Mismatches 22; Indels 2; Gaps 2;

QY 1 MSPDRPFKORSFADCKEVOQIRDOHPSKIPVIERYKGEKQLPVLDTKFLVDPDHYNM 59
    |||:|||||:| |::|:|||||:|||||:|||||:|||||:|||||:|||||:
DB 32 MPSEKTFKQRTFEQRYVDVRLIREQHTKIPVIERKKGKQLPVLDTKFLVDPDHYNM 91
QY 60 MSELVKIIRRLQLNPTQAFLLVNOHSMVSVSTPIADIYEOEKDEDEGFLYMYVASOET 118
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 92 MSELIRIIRRLQLNANOAFLLVNGHSMVSVSTPISEYSEKEDDEGFLYMYLCPPPE 151
QY 119 FG 120
    :|
DB 152 YG 153

RESULT 4
US-10-293-244-3825
;; Sequence 3825, Application US/10293244
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq, Inc.
;; APPLICANT: Tang, Y. Tom et al
;; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;; FILE REFERENCE: 21272-029
;; CURRENT APPLICATION NUMBER: US/10/293,244
;; CURRENT FILING DATE: 2002-11-12
;; PRIOR APPLICATION NUMBER: Not Yet Assigned
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: 09/728,422
;; PRIOR FILING DATE: 2000-11-30
;; PRIOR APPLICATION NUMBER: 09/693,325
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 09/663,561
;; PRIOR FILING DATE: 2000-09-15
;; PRIOR APPLICATION NUMBER: 09/654,936
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 09/620,325
;; PRIOR FILING DATE: 2000-07-19
;; PRIOR APPLICATION NUMBER: 09/598,075
;; PRIOR FILING DATE: 2000-06-20
;; PRIOR APPLICATION NUMBER: 09/560,875
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 09/496,914
;; PRIOR FILING DATE: 2000-02-03
;; NUMBER OF SEQ ID NOS: 3960
;; SOFTWARE: Custom
;; SEQ ID NO: 3825
;; LENGTH: 120
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-293-244-3825

Query Match      48.2%; Score 301; DB 6; Length 120;
Best Local Similarity 59.2%; Pred. No. 8.2e-27;
Matches 58; Conservative 16; Mismatches 24; Indels 0; Gaps 0;

QY 23 IRDOHPSKIPVIERYKGEKQLPVLDTKFLVDPDHYNMSELVKIIRRLQLNPTQAFLL 82
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2 IRAKFPNKIPVYVERPETFLLPPLDKTKFLVPOELTWTOTLIRSRMVLRAVEAFYLL 61
QY 83 VNQHSMSVSTPIADIYEOEKDEDEGFLYMYVASOETFG 120
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 62 VNKSLSVMSATMAEIRDRYKDEDEGFLYMYTASOETFG 99

RESULT 5
US-10-293-244-1857
;; Sequence 1857, Application US/10293244
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq, Inc.
;; APPLICANT: Tang, Y. Tom et al
```


Query Match	43.58;	Score 272;	DB 6;	Length 135;
Best Local Similarity	55.68;	Pred. No. 1,76-23;		
Matches	55;	Conservative 18;	Mismatches 26;	Indels 0;
			Gaps	
QY	1	MPSDPFRKORRSADSCKEVOQIIRDDHSPKPIPIERKYGKGLPDLDTKTLFVPHVM	60	
Db	7	IPSVPRPKOKSLAIKOEYAGIRAKFPAKPIIPVWVPRPREFPLPLDKTKFLVDELTM	66	
QY	61	SELVKIIRRRQLNPTQAPFLVNOHSMVSVSTPIADIV	99	
	67	TOFLSIIRSRMVLRAATEAFYLLVNNKSLVMSMTAKETI	105	

APPLICANT: Zhang, Bing
APPLICANT: Gibson, Bradford W

RESULT 8
US-10-408-765A-1021
; Sequence 1021 Application US/10408765A
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.

```

APPLICANT: AZIMZAI, Valda; GIERZEN, Kimberly J.;
APPLICANT: TANG, Y. Tom; WARREN, Bridget A.;
APPLICANT: MASON, Patricia M.; BUREFORD, Neil;
APPLICANT: HAFALIA, April J.A.; LEE, Ernestine A.;
APPLICANT: YANG, Junming; GORVAD, Ann E.;
APPLICANT: EMERLING, Brooke M.; MARQUIS, Joseph P.;
APPLICANT: LEE, Soo Yeun; SMARNAKAR, Anita;
APPLICANT: REDDY, Roopa M.; JIANG, Xin;
APPLICANT: JACKSON, Alan A.
TITLE OF INVENTION: NUCLEIC ACID-ASSOCIATED PROTEINS
FILE REFERENCE: PF-0869 USN
CURRENT APPLICATION NUMBER: US/10/451,901
CURRENT FILING DATE: 2003-06-23
PRIORITY APPLICATION NUMBER: PCT/US01/50256
PRIORITY FILING DATE: 2001-12-19
PRIORITY APPLICATION NUMBER: US 60/257,714
PRIORITY FILING DATE: 2000-12-21
PRIORITY APPLICATION NUMBER: US 60/260,081
PRIORITY FILING DATE: 2001-01-05
PRIORITY APPLICATION NUMBER: US 60/262,302
PRIORITY FILING DATE: 2001-01-16
PRIORITY APPLICATION NUMBER: US 60/263,823
PRIORITY FILING DATE: 2001-01-23
PRIORITY APPLICATION NUMBER: US 60/266,088
PRIORITY FILING DATE: 2001-02-02
PRIORITY APPLICATION NUMBER: US 60/348,442
PRIORITY FILING DATE: 2001-10-29
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PERL Program
SEQ ID NO 4
LENGTH: 602
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No: 7472664CD1
US-10-451-901-4

Query Match 11.0%; Score 68.5; DB 6; Length 602;
Best Local Similarity 29.3%; Pred. No. 7;
Matches 24; Conservative 12; Mismatches 37; Indels 9; Gaps 3;

QY 41 EKQLPVLDK-TKELVDPDHVMSLELYIKRRRLQNPQAFLLVNOHSMVSVSTPIADIY 99
Db 525 EKALAIISQTPSLVYDHETRLKILKIVVKKSQ---NYNIFOLENLVAIVISQC-----IY 576
QY 100 EOEKDEDEGLVYVYASQETGFC 121
Db 577 RHKKDHDKTSLIQKMEDEVENF 598

RESULT 11
US-10-603-113-14286
: Sequence 14286; Application US/10603113
GENERAL INFORMATION:
APPLICANT: Keith Weinstein;
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBIDUS
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/10/603,113
CURRENT FILING DATE: 2003-06-24
PRIORITY APPLICATION NUMBER: US/09/248,796
PRIORITY FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 28206
SEQ ID NO 14286
LENGTH: 1648
TYPE: PRT
ORGANISM: Candida albicans

Query Match 11.0%; Score 68.5; DB 6; Length 1648;
Best Local Similarity 20.3%; Pred. No. 26;
Matches 29; Conservative 31; Mismatches 50; Indels 33; Gaps 6;

```

[illegible]

Matches 31; Conservative 22; Mismatches 39; Indels 31; Gaps 8;

QY 16 RCKEVQOI-RDQHPSPKIPVIERYKGEK-----QLPVLDTKFLV---PDHVNMSSELVKI 66

Db 159 RSHQVEKLENEKEFINIPRISSPSNGEKPFTTDLP--DQLEDLIVASPDINCCTIILT 216

QY 67 IRRR---LQLNPTQAF---LLVNQHSMSVSTPIADIYEQEKDEDEGFLMYVASQET 118

Db 217 IHDPPEIWANINPTNGVIEIETNSTMWLEHPLVQ---IMDKIQKEKDP-----NQEK 264

QY	119	FGF	121
		11.	

Db 265 FGY 267

RESULT 15
HE-10-603

; Sequence 17078, Application US/10603113

APPLICANT: Keith Weinstock et al

1. TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

REFERENCE: 10/150-132
CURRENT APPLICATION NUMBER: US/10/603,113

CORRENT FILING DATE: 2003-08-24
 PRIOR APPLICATION NUMBER: US/09/248,796

NUMBER OF SEQ ID NOS: 28206

```

; SEQ ID NO 1/0/8
; LENGTH: 454

```

```

; TYPE: PRT
; ORGANISM: Candida albicans

```

US-10-603-113-17078

Query Match	10.6%	Score 66
Best Local Similarity	22.7%	Pred NC

Matches 25; Conservative

QY 8 KÖRSEADCKEVÖÖT

Db 233 K

QY 68 R

Db 283

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

Job time : 21 secs

Query Match	Score	DB	Length
10.68;	66;	6;	454;
32.78;	66;	6;	454;
32.78;	66;	6;	454;

Matches 25; Conservative 21; Mismatches 54; Indels 10; Gaps 1;

8 KÖRSEFADCKEVÖQIRDÖHPSKIPVIERYKGEKÖLPVLDKTKFLVPDHNMSLVKII 67

Db 233 KYPKEFASRLK-----NLSYRYPGVGESYLDVLTRLRPLIAEVERTDHLII 283

QY 68 RRLQLNPTQAEFLLVNQHSMVSVSTPIADIYEQEKDEDEGFLMYVASQE 117

Db 283 SHRVSRIILAYFLNLDKSAIGELDVPLHTLYCLELHPYGTDTMYEYDE 332

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

om nucleic - nucleic search, using sw model

Run on: August 15, 2003, 08:09:17 : Search time 2934 Seconds
(without alignments)
7147.724 Million cell updates/sec

Title: US-09-904-603-2
Perfect score: 640
Sequence: 1 CTCGCCGAGCGCGCGCC.....GCCGCCCTAGTCATACAGGCA 640

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 32997241 seqs, 1638322548 residues

Total number of hits satisfying chosen parameters: 65944482

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1: /cgn2_6/prodata/1/pna/PCRUS_COMB.seq:*
2: /cgn2_6/prodata/1/pna/PCRUS_COMB.seq.OLD:*
3: /cgn2_6/prodata/1/pna/US06_COMB.seq:*
4: /cgn2_6/prodata/1/pna/US07_COMB.seq:*
5: /cgn2_6/prodata/1/pna/US080_COMB.seq:*
6: /cgn2_6/prodata/1/pna/US081_COMB.seq:*
7: /cgn2_6/prodata/1/pna/US082_COMB.seq:*
8: /cgn2_6/prodata/1/pna/US083_COMB.seq:*
9: /cgn2_6/prodata/1/pna/US084_COMB.seq:*
10: /cgn2_6/prodata/1/pna/US085_COMB.seq:*
11: /cgn2_6/prodata/1/pna/US086_COMB.seq:*
12: /cgn2_6/prodata/1/pna/US087_COMB.seq:*
13: /cgn2_6/prodata/1/pna/US088_COMB.seq:*
14: /cgn2_6/prodata/1/pna/US089_COMB.seq:*
15: /cgn2_6/prodata/1/pna/US090_COMB.seq:*
16: /cgn2_6/prodata/1/pna/US091_COMB.seq:*
17: /cgn2_6/prodata/1/pna/US092A_COMB.seq:*
18: /cgn2_6/prodata/1/pna/US092B_COMB.seq:*
19: /cgn2_6/prodata/1/pna/US093A_COMB.seq:*
20: /cgn2_6/prodata/1/pna/US093B_COMB.seq:*
21: /cgn2_6/prodata/1/pna/US094_COMB.seq:*
22: /cgn2_6/prodata/1/pna/US095A_COMB.seq:*
23: /cgn2_6/prodata/1/pna/US095B_COMB.seq:*
24: /cgn2_6/prodata/1/pna/US095C_COMB.seq:*
25: /cgn2_6/prodata/1/pna/US095D_COMB.seq:*
26: /cgn2_6/prodata/1/pna/US096A_COMB.seq:*
27: /cgn2_6/prodata/1/pna/US096B_COMB.seq:*
28: /cgn2_6/prodata/1/pna/US096C_COMB.seq:*
29: /cgn2_6/prodata/1/pna/US096D_COMB.seq:*
30: /cgn2_6/prodata/1/pna/US096E_COMB.seq:*
31: /cgn2_6/prodata/1/pna/US097A_COMB.seq:*
32: /cgn2_6/prodata/1/pna/US097B_COMB.seq:*
33: /cgn2_6/prodata/1/pna/US097C_COMB.seq:*
34: /cgn2_6/prodata/1/pna/US098A_COMB.seq:*
35: /cgn2_6/prodata/1/pna/US098B_COMB.seq:*
36: /cgn2_6/prodata/1/pna/US098C_COMB.seq:*
37: /cgn2_6/prodata/1/pna/US098D_COMB.seq:*
38: /cgn2_6/prodata/1/pna/US099A_COMB.seq:*
39: /cgn2_6/prodata/1/pna/US099B_COMB.seq:*
40: /cgn2_6/prodata/1/pna/US099C_COMB.seq:*
41: /cgn2_6/prodata/1/pna/US099D_COMB.seq:*
42: /cgn2_6/prodata/1/pna/US099E_COMB.seq:*
43: /cgn2_6/prodata/1/pna/US099F_COMB.seq:*

Result		No.		Score		Match		Query		Length (bp)		Description	
1	640	100.0	640	46	US-09-904-603-2	Sequence 2, Aff:							
2	627.4	98.0	925	25	US-09-572-411-5453	Sequence 5453, Aff:							
3	627.4	98.0	945	47	US-10-170-245-27458	Sequence 27458, Aff:							
4	627.4	98.0	977	21	US-09-471-275-1555	Sequence 1555, Aff:							

44: /cgn2_6/prodata/1/pna/US100A_COMB.seq:*
45: /cgn2_6/prodata/1/pna/US100B_COMB.seq:*
46: /cgn2_6/prodata/1/pna/US101A_COMB.seq:*
47: /cgn2_6/prodata/1/pna/US101B_COMB.seq:*
48: /cgn2_6/prodata/1/pna/US102A_COMB.seq:*
49: /cgn2_6/prodata/1/pna/US102B_COMB.seq:*
50: /cgn2_6/prodata/1/pna/US103A_COMB.seq:*
51: /cgn2_6/prodata/1/pna/US103B_COMB.seq:*
52: /cgn2_6/prodata/1/pna/US104A_COMB.seq:*
53: /cgn2_6/prodata/1/pna/US104B_COMB.seq:*
54: /cgn2_6/prodata/1/pna/US6000_COMB.seq:*
55: /cgn2_6/prodata/1/pna/US6001_COMB.seq:*
56: /cgn2_6/prodata/1/pna/US6002_COMB.seq:*
57: /cgn2_6/prodata/1/pna/US6003_COMB.seq:*
58: /cgn2_6/prodata/1/pna/US6004_COMB.seq:*
59: /cgn2_6/prodata/1/pna/US6005_COMB.seq:*
60: /cgn2_6/prodata/1/pna/US6006_COMB.seq:*
61: /cgn2_6/prodata/1/pna/US6007_COMB.seq:*
62: /cgn2_6/prodata/1/pna/US6008_COMB.seq:*
63: /cgn2_6/prodata/1/pna/US6009_COMB.seq:*
64: /cgn2_6/prodata/1/pna/US6010_COMB.seq:*
65: /cgn2_6/prodata/1/pna/US6011_COMB.seq:*
66: /cgn2_6/prodata/1/pna/US6012_COMB.seq:*
67: /cgn2_6/prodata/1/pna/US6013_COMB.seq:*
68: /cgn2_6/prodata/1/pna/US6014_COMB.seq:*
69: /cgn2_6/prodata/1/pna/US6015_COMB.seq:*
70: /cgn2_6/prodata/1/pna/US6016_COMB.seq:*
71: /cgn2_6/prodata/1/pna/US6017_COMB.seq:*
72: /cgn2_6/prodata/1/pna/US6018_COMB.seq:*
73: /cgn2_6/prodata/1/pna/US6019_COMB.seq:*
74: /cgn2_6/prodata/1/pna/US6020_COMB.seq:*
75: /cgn2_6/prodata/1/pna/US6021_COMB.seq:*
76: /cgn2_6/prodata/1/pna/US6022_COMB.seq:*
77: /cgn2_6/prodata/1/pna/US6023A_COMB.seq:*
78: /cgn2_6/prodata/1/pna/US6023B_COMB.seq:*
79: /cgn2_6/prodata/1/pna/US6024_COMB.seq:*
80: /cgn2_6/prodata/1/pna/US6025_COMB.seq:*
81: /cgn2_6/prodata/1/pna/US6026_COMB.seq:*
82: /cgn2_6/prodata/1/pna/US6027_COMB.seq:*
83: /cgn2_6/prodata/1/pna/US6028_COMB.seq:*
84: /cgn2_6/prodata/1/pna/US6029_COMB.seq:*
85: /cgn2_6/prodata/1/pna/US6030_COMB.seq:*
86: /cgn2_6/prodata/1/pna/US6031_COMB.seq:*
87: /cgn2_6/prodata/1/pna/US6032_COMB.seq:*
88: /cgn2_6/prodata/1/pna/US6033_COMB.seq:*
89: /cgn2_6/prodata/1/pna/US6034_COMB.seq:*
90: /cgn2_6/prodata/1/pna/US6035_COMB.seq:*
91: /cgn2_6/prodata/1/pna/US6036_COMB.seq:*
92: /cgn2_6/prodata/1/pna/US6037_COMB.seq:*
93: /cgn2_6/prodata/1/pna/US6038_COMB.seq:*
94: /cgn2_6/prodata/1/pna/US6039_COMB.seq:*
95: /cgn2_6/prodata/1/pna/US6040_COMB.seq:*
96: /cgn2_6/prodata/1/pna/US6041_COMB.seq:*
97: /cgn2_6/prodata/1/pna/US6042_COMB.seq:*
98: /cgn2_6/prodata/1/pna/US6043_COMB.seq:*
99: /cgn2_6/prodata/1/pna/US6044_COMB.seq:*
100: /cgn2_6/prodata/1/pna/US6045_COMB.seq:*
101: /cgn2_6/prodata/1/pna/US6046_COMB.seq:*
102: /cgn2_6/prodata/1/pna/US6047_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

5 627.4 98.0 977 21 US-09-488-725B-8516 Sequence 8516, App
6 627.4 98.0 1000 1 PCT-US01-18569-361 Sequence 361, App
7 627.4 98.0 1000 2 PCT-US01-18569-361 Sequence 361, App
8 627.4 98.0 1000 49 US-10-264-049-361 Sequence 361, App
9 627.4 98.0 1009 77 US-60-234-690-7194 Sequence 7194, App
10 627.4 98.0 1009 87 US-60-323-843-7194 Sequence 7194, App
11 627.4 98.0 1019 52 US-10-405-027-30 Sequence 30, App
12 627.4 98.0 1356 21 US-09-428-151A-9451 Sequence 9451, App
13 627 98.0 951 1 PCT-US98-15949-94 Sequence 94, App
14 627 98.0 951 1 PCT-US98-15949-94 Sequence 94, App
15 605.4 94.6 1017 71 US-60-172-373-10400 Sequence 10400, A
16 593 92.7 719 28 US-09-644-871-273 Sequence 273, App
17 507 79.2 539 20 US-09-354-899-5690 Sequence 5690, App
18 497 79.2 539 21 US-09-428-151A-5852 Sequence 5852, App
19 497 77.7 557 20 US-09-397-761A-2899 Sequence 2899, App
20 497 77.7 557 20 US-09-838-601-2899 Sequence 2899, App
21 493 77.0 505 20 US-09-371-168-264 Sequence 264, App
22 493 77.0 505 34 US-09-801-833-264 Sequence 264, App
23 482.2 75.3 605 80 US-60-253-652-16982 Sequence 16982, A
24 480.4 75.1 600 80 US-60-253-652-14610 Sequence 14610, A
25 475.2 72.3 570 28 US-09-652-109-5783 Sequence 5783, App
26 462.6 72.3 557 23 US-10-405-027-1218 Sequence 1218, App
27 461.2 72.1 557 23 US-09-534-854-14761 Sequence 14761, A
28 460.4 71.9 966 82 US-10-170-235-26846 Sequence 26846, A
29 460.4 71.9 1014 17 US-09-236-557-26 Sequence 12073, A
30 460.4 71.9 1014 17 US-09-236-557-26 Sequence 26, App
31 460.4 71.3 704 1 PCT-US01-08631-9784 Sequence 9784, App
32 456.4 71.3 704 2 PCT-US01-08631-9784 Sequence 3509, App
33 456.4 71.2 516 28 US-09-644-871-3509 Sequence 14460, A
34 454.4 70.7 792 19 US-09-205-070-14460 Sequence 14460, A
35 452.4 70.7 792 19 US-09-340-623-14460 Sequence 867, App
36 452.4 70.7 792 20 US-09-359-922-867 Sequence 867, App
37 452.4 70.7 792 20 US-09-359-922-867 Sequence 14460, A
38 452.4 70.7 792 37 US-09-898-888-14460 Sequence 14460, A
39 452.4 70.7 792 37 US-09-898-888-14460 Sequence 867, App
40 452.4 70.7 792 38 US-09-919-002-867 Sequence 867, App
41 450.6 70.4 470 22 US-09-528-409-60388 Sequence 60388, A
42 450.6 70.4 470 39 US-09-933-524-60388 Sequence 60388, A
43 450.6 70.4 470 39 US-09-933-524-60388 Sequence 60388, A
44 450.6 70.4 470 39 US-09-933-524-60388 Sequence 682, App
45 449.8 70.3 516 25 US-09-572-411-682 Sequence 682, App

ALIGNMENTS

RESULT 1
US-09-904-603-2
Sequence 2, Application US/09904603
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
Goli, Surya K.
TITLE OF INVENTION: NOVEL MICROTUBULE-ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Inocyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/904,603
FILING DATE: 12-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/805,117
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0211 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 640 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PHYNOT03
CLONE: 1441378
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-904-603-2
Query Match 100.0%; Score 640; DB 38; Length 640;
Best Local Similarity 100.0%; Pred. No. 1.6e-109; Mismatches 0; Gaps 0;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0:
1 CTCCGCGACCGCGCCGCTGCTCAGCGCGAGCCCGGAGCCCTTGAGCGCGAGCGC 60
1 CTCCGCGACCGCGCGCCGCTGCTCAGCGCGAGCCCGGAGCCCTTGAGCGCGAGCGC 60
61 GGACCG 120
61 GGACCG 120
121 GCCGAGCG 180
121 GCCGAGCG 180
181 CGACCGCGTGAAGAGGATGACAGAGATCCGCGCGCGCGCGCGCGCGCGCGCGCG 240
181 CGACCGCGTGAAGAGGATGACAGAGATCCGCGCGCGCGCGCGCGCGCGCGCGCG 240
241 CATGAGCGCGTGAAGAGGATGACAGAGATCCGCGCGCGCGCGCGCGCGCGCGCGCG 300
241 CATGAGCGCGTGAAGAGGATGACAGAGATCCGCGCGCGCGCGCGCGCGCGCGCGCG 300
301 CCCGACCGATGTAACATGAGCGAGTGTGATCATCGCGCGCGCGCGCGCGCGCGCG 360
301 CCCGACCGATGTAACATGAGCGAGTGTGATCATCGCGCGCGCGCGCGCGCGCGCG 360
361 CCCGACCGAGCGCTTCTCTGCTGCTGTAACACGACGATGAGTGTGTCACGCGC 420
361 CCCGACCGAGCGCTTCTCTGCTGCTGTAACACGACGATGAGTGTGTCACGCGC 420
421 CATGCGGACATGTAACATGAGCGAGTGTGATCATCGCGCGCGCGCGCGCGCGCG 480
421 CATGCGGACATGTAACATGAGCGAGTGTGATCATCGCGCGCGCGCGCGCGCGCG 480
481 CTCCGAGAAACCTTGGCTTCTGAGCGAGAGTGTGATCATCGCGCGCGCGCGCGCG 540
481 CTCCGAGAAACCTTGGCTTCTGAGCGAGAGTGTGATCATCGCGCGCGCGCGCGCG 540
541 GCGCGCGGTGAGCG 600
541 GCGCGCGGTGAGCG 600
601 GGTGCGGTGAGCGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 640
601 GGTGCGGTGAGCGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 640

RESULT 2
US-09-572-411-5453
Sequence 5453, Application US/09572411
GENERAL INFORMATION:

NAME/KEY: misc.feature
LOCATION: (155)...(517)
OTHER INFORMATION: similar to q1455109 in the genepept database release 114,
OTHER INFORMATION: Run with FASTX 3.3t00, default parameters
US-09-488-725B-8516

Query Match 98.0%: Score 627.4; DB 21; Length 977;
Best Local Similarity 99.7%: Pred. No. 3,4e-107;
Matches 639; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
QY 1 CTCCCGCAGCCGCGAGCGCGCTGCTCAGCGCGAGCCCGCGAGCCCTTGAGCGCGAGCGCG 60
DB 15 CTCCCGCAGCCGCGAGCGCGCGCTGCTCAGCGCGAGCGCGCGAGCCCTTGAGCGCGAGCGCG 74
QY 61 GGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 119
DB 75 GGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 134
QY 120 CGCCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
DB 145 CGCCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 194
QY 180 CGCAGCGCGGTGAAGAGGTACAGCAGATCCCGGACCGACCGACCGACCGACCGACCGACCG 239
DB 195 CGCAGCGCGGTGAAGAGGTACAGCAGATCCCGGACCGACCGACCGACCGACCGACCGACCG 254
QY 240 TCATCGAGCGGTACAGAGGTAGAGAGCGCTGCCGCTGCTGAGCAGAGCGCAAGTTTGG 299
DB 255 TCATCGAGCGGTACAGAGGTAGAGAGCGCTGCCGCTGCTGAGCAGAGCGCAAGTTTGG 314
QY 300 TCCCGAGCGCATGTACATAGAGAGGTGTCAGATGATCCGGCGCGCGCTGAGAGTGA 359
DB 315 TCCCGAGCGCATGTACATAGAGAGGTGTCAGATGATCCGGCGCGCGCTGAGAGTGA 374
QY 360 ACCCGAGCGAGCGCTTCTCTGCTGCTGTAACACAGCAGCAGCAGTGGTGGTGTGTCACGC 419
DB 375 ACCCGAGCGAGCGCTTCTCTGCTGCTGTAACACAGCAGCAGCAGTGGTGGTGTGTCACGC 434
QY 420 CCATCGCGGAGCATGTACAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 479
DB 435 CCATCGCGGAGCATGTACAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 494
QY 480 CCTCCGAGAGAACTTGGGCTTCTGAGCGCAGCAGTAGGGGGGCTGGGCTGGAGTCCGG 539
DB 495 CCTCCGAGAGAACTTGGGCTTCTGAGCGCAGCAGTAGGGGGGCTGGGCTGGAGTCCGG 554
QY 540 GGGCGCGCGGTAGAGCGCTGCGCCAGAGAGCTTCTGCTGTAAGCTGAGCTGCTCTACCG 599
DB 555 GGGCGCGCGGTAGAGCGCTGCGCCAGAGAGCTTCTGCTGTAAGCTGAGCTGCTCTACCG 614
QY 600 TGGTGGGTGGCGAGCGATGTGCCCGCTAGTCAAGAGGCA 640
DB 615 TGGTGGGTGGCGAGCGATGTGCCCGCTAGTCAAGAGGCA 655
```

RESULT 6
PCT-US01-18569-361
Sequence 461, Application PC/TUS0118569

GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA133PCT
CURRENT APPLICATION NUMBER: PCT/US01/18569
CURRENT FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: 60/209,467
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 4360
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 361
LENGTH: 1000
TYPE: DNA
ORGANISM: Homo sapiens
PCT-US01-18569-361

Query Match 98.0%: Score 627.4; DB 1; Length 1000;
Best Local Similarity 99.7%: Pred. No. 3,4e-107;
Matches 639; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
QY 1 CTCCCGTAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
DB 6 CTCCCGTAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 65
QY 61 GGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 119
DB 66 GGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 125
QY 120 CGCCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
DB 125 CGCCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 185
QY 180 CGCAGCGCGGTGAAGAGGTACAGCAGATCCCGGACCGACCGACCGACCGACCGACCGACCG 239
DB 186 CGCAGCGCGGTGAAGAGGTACAGCAGATCCCGGACCGACCGACCGACCGACCGACCGACCG 245
QY 240 TCATCGAGCGGTACAGAGGTAGAGAGCGCTGCCGCTGCTGAGCAGAGCGCAAGTTTGG 299
DB 246 TCATCGAGCGGTACAGAGGTAGAGAGCGCTGCCGCTGCTGAGCAGAGCGCAAGTTTGG 314
QY 300 TCCCGAGCGCATGTACATAGAGAGGTGTCAGATGATCCGGCGCGCGCTGAGAGTGA 359
DB 306 TCCCGAGCGCATGTACATAGAGAGGTGTCAGATGATCCGGCGCGCGCTGAGAGTGA 374
QY 360 ACCCGAGCGAGCGCTTCTCTGCTGCTGTAACACAGCAGCAGCAGTGGTGGTGTGTCACGC 419
DB 366 ACCCGAGCGAGCGCTTCTCTGCTGCTGTAACACAGCAGCAGCAGTGGTGGTGTGTCACGC 425
QY 420 CCATCGCGGAGCATGTACAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 479
DB 426 CCATCGCGGAGCATGTACAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 485
QY 480 CCTCCGAGAGAACTTGGGCTTCTGAGCGCAGCAGTAGGGGGGCTGGGCTGGAGTCCGG 539
DB 486 CCTCCGAGAGAACTTGGGCTTCTGAGCGCAGCAGTAGGGGGGCTGGGCTGGAGTCCGG 545
QY 540 GGGCGCGCGGTAGAGCGCTGCGCCAGAGAGCTTCTGCTGTAAGCTGAGCTGCTCTACCG 599
DB 546 GGGCGCGCGGTAGAGCGCTGCGCCAGAGAGCTTCTGCTGTAAGCTGAGCTGCTCTACCG 605
QY 600 TGGTGGGTGGCGAGCGATGTGCCCGCTAGTCAAGAGGCA 640
DB 606 TGGTGGGTGGCGAGCGATGTGCCCGCTAGTCAAGAGGCA 646
```

RESULT 7
PCT-US01-18569-361
Sequence 361, Application PC/TUS0118569

GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA133PCT
CURRENT APPLICATION NUMBER: PCT/US01/18569
CURRENT FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: 60/209,467
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 4360
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 361
LENGTH: 1000
TYPE: DNA
ORGANISM: Homo sapiens
PCT-US01-18569-361

Query Match 98.0%: Score 627.4; DB 2; Length 1000;
Best Local Similarity 99.7%: Pred. No. 3,4e-107;
Matches 639; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

OY 1 CTCCGCGAGCGCGCGCGCTGCTCAGCGCGAGCGCCCGGAGCCCTTGAGCGCGAGCGGC 60
DB 6 CTCCGCGAGCGCGCGCGCGCTGCTCAGCGCGAGCGCCCGGAGCCCTTGAGCGCGAGCGGC 65
OY 61 GGAG-CCCCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 119
DB 66 GGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 125
OY 120 CGCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
DB 126 CGCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 185
OY 180 CCGACCGCTGTAAAGAGGTACAGCAGATCCGCGACAGCAGCCCGCAGCAAAATCCCGGTGA 239
DB 186 CCGACCGCTGTAAAGAGGTACAGCAGATCCGCGACAGCAGCCCGCAGCAAAATCCCGGTGA 245
OY 240 TCATGACGCGCTACAGAGGTGAGAGCAGCTGCGCGCTGCGACAGCAAGTTTGG 299
DB 246 TCATGACGCGCTACAGAGGTGAGAGCAGCTGCGCGCTGCGACAGCAAGTTTGG 305
OY 300 TCCCGCGACCATGTCAACATGAGCGAGTTGTCAGATCAGCGCGCGCGCGCGCGCGCGCG 359
DB 306 TCCCGCGACCATGTCAACATGAGCGAGTTGTCAGATCAGCGCGCGCGCGCGCGCGCGCG 365
OY 360 ACCCGAGCGAGCGCTTCTTCTGCTGCTGTAACAGCAGCAGATGCTGCTGCGCGCGCG 419
DB 366 ACCCGAGCGAGCGCTTCTTCTGCTGCTGTAACAGCAGCAGATGCTGCTGCGCGCGCG 425
OY 420 CCATGCGCGACATGTACAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 479
DB 426 CCATGCGCGACATGTACAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 485
OY 480 CCTCCAGAGAAACCTTGGCTTGTGAGCGAGCAGTGAAGGGGCGCTGCGCGCGCGCGCG 539
DB 486 CCTCCAGAGAAACCTTGGCTTGTGAGCGAGCAGTGAAGGGGCGCTGCGCGCGCGCGCG 545
OY 540 GGGCGCGCGTGAAGCGCTGCGCGCGAGAGAGCTTCTGCTGTAACAGCAGCAGTGAAG 599
DB 546 GGGCGCGCGTGAAGCGCTGCGCGCGAGAGAGCTTCTGCTGTAACAGCAGCAGTGAAG 605
OY 600 TGTGCGCTGGGCGAGCATGTGCCCGCTTACTAGCAGAGGCA 640
DB 606 TGTGCGCTGGGCGAGCATGTGCCCGCTTACTAGCAGAGGCA 646

```

```

RESULT 8
US-10-264-049-361
: Sequence 361, Application US/10264049
: GENERAL INFORMATION:
: APPLICANT: Birse et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PA13PI
: CURRENT APPLICATION NUMBER: US/10/264,049
: CURRENT FILING DATE: 2002-10-04
: PRIOR APPLICATION NUMBER: PCT/US01/18569
: PRIOR FILING DATE: 2001-06-07
: PRIOR APPLICATION NUMBER: US 60/209,467
: PRIOR FILING DATE: 2000-06-07
: NUMBER OF SEQ ID NOS: 4360
: SOFTWARE: PatentIn Ver. 3.1
: SEQ ID NO 361
: LENGTH: 1000
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-264-049-361

```

```

Query Match 98.0%, Score 627.4, DB 49, Length 1000,
Best Local Similarity 99.7%, Pred. No. 3.4e-107,
Matches 639, Conservative 0, Mismatches 1, Indels 1, Gaps 1:
OY 1 CTCCGCGAGCGCGCGCGCGCTGCTCAGCGCGAGCGCCCGGAGCCCTTGAGCGCGAGCGGC 60
DB 6 CTCCGCGAGCGCGCGCGCGCGCTGCTCAGCGCGAGCGCCCGGAGCCCTTGAGCGCGAGCGGC 65

```

```

OY 61 GGAG-CCCCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 119
DB 66 GGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 125
OY 120 CGCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
DB 126 CGCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 185
OY 180 CCGACCGCTGTAAAGAGGTACAGCAGATCCGCGAGCAGCAGCCCGCAGCAAAATCCCGGTGA 239
DB 186 CCGACCGCTGTAAAGAGGTACAGCAGATCCGCGAGCAGCAGCCCGCAGCAAAATCCCGGTGA 245
OY 240 TCATGACGCGCTACAGAGGTGAGAGCAGCTGCGCGCTGCGACAGCAAGTTTGG 299
DB 246 TCATGACGCGCTACAGAGGTGAGAGCAGCTGCGCGCTGCGACAGCAAGTTTGG 305
OY 300 TCCCGCGACCATGTCAACATGAGCGAGTTGTCAGATCAGCGCGCGCGCGCGCGCGCGCG 359
DB 306 TCCCGCGACCATGTCAACATGAGCGAGTTGTCAGATCAGCGCGCGCGCGCGCGCGCGCG 365
OY 360 ACCCGAGCGAGCGCTTCTTCTGCTGCTGTAACAGCAGCAGATGCTGCTGCGCGCGCG 419
DB 366 ACCCGAGCGAGCGCTTCTTCTGCTGCTGTAACAGCAGCAGATGCTGCTGCGCGCGCG 425
OY 420 CCATGCGCGACATGTACAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 479
DB 426 CCATGCGCGACATGTACAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 485
OY 480 CCTCCAGAGAAACCTTGGCTTGTGAGCGAGCAGTGAAGGGGCGCTGCGCGCGCGCGCG 539
DB 486 CCTCCAGAGAAACCTTGGCTTGTGAGCGAGCAGTGAAGGGGCGCTGCGCGCGCGCGCG 545
OY 540 GGGCGCGCGTGAAGCGCTGCGCGCGAGAGAGCTTCTGCTGTAACAGCAGCAGTGAAG 599
DB 546 GGGCGCGCGTGAAGCGCTGCGCGCGAGAGAGCTTCTGCTGTAACAGCAGCAGTGAAG 605
OY 600 TGTGCGCTGGGCGAGCATGTGCCCGCTTACTAGCAGAGGCA 640
DB 606 TGTGCGCTGGGCGAGCATGTGCCCGCTTACTAGCAGAGGCA 646

```

```

RESULT 9
US-60-234-690-7194
: Sequence 7194, Application US/60234690
: GENERAL INFORMATION:
: APPLICANT: Lal, Preeti
: APPLICANT: Diep, Dinh
: TITLE OF INVENTION: METHOD FOR THE IDENTIFICATION OF SEQUENCE POLYMORPHISMS USING
: FILE REFERENCE: POLYNUCLEOTIDE SEQUENCE DATABASES, AND SINGLE NUCLEOTIDE
: CURRENT APPLICATION NUMBER: US/60/234,690
: CURRENT FILING DATE: 2000-09-21
: NUMBER OF SEQ ID NOS: 12060
: SOFTWARE: PERL Program
: SEQ ID NO 7194
: LENGTH: 1009
: TYPE: DNA
: ORGANISM: Homo sapiens
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 247776.13
US-60-234-690-7194

```

```

Query Match 98.0%, Score 627.4, DB 77, Length 1009,
Best Local Similarity 99.7%, Pred. No. 3.4e-107,
Matches 639, Conservative 0, Mismatches 1, Indels 1, Gaps 1:
OY 1 CTCCGCGAGCGCGCGCGCGCGCTGCTCAGCGCGAGCGCCCGGAGCCCTTGAGCGCGAGCGGC 60
DB 57 CTCCGCGAGCGCGCGCGCGCGCGCTGCTCAGCGCGAGCGCCCGGAGCCCTTGAGCGCGAGCGGC 116

```

```

UY      61  TGGAG - CCCCAGAGCCCCCAACCCGAGACACATCCGCGCGCCGACAGAGCCGCGCTGGG  119
Db      117  GAGAGCCCCCGAGCCCCCAACCCGAGACACATCCCGCGCCGCAAGCCGCCGCGCTGGG  176
UY      120  CGCCGAGCGCGGCGCGCGGATGCCCTCGAGCCCGGCTTTCAAGCAGCGCGGAGCTTGG  179
Db      177  CGCCGAGCGCGGCGCGCGGATGCCCTCGAGCCCGGCTTTCAAGCAGCGCGGAGCTTGG  236
UY      180  CGGACCGGCTGTAAAGSAGGTACAGCAGATCCCGGACGACAGCCCCAGCAAAATCCCGGTA  239
Db      237  CGGACCGGCTGTAAAGSAGGTACAGCAGATCCCGGACGACAGCCCCAGCAAAATCCCGGTA  296
UY      240  TCATGAGCGGCTACAAAGGSGTAGAAGCAGCTGCCGCTCTGGGACAAGACCAAGTTTGG  299
Db      297  TCATGAGCGGCTACAAAGGSGTAGAAGCAGCTGCCGCTCTGGGACAAGACCAAGTTTGG  356
UY      357  TCCCGGACCATATGTCAACATGAGCGGAGTTGGTCAAGATCATCCGCGCGCCCTGACCTGA  416
Db      417  ACCCGACGAGGCTTCTTCCTGCTGGTGAACCCAGACAGCATGGTGAATGTGTACAGCG  476
UY      420  CCATCGCGACATCTACGAGCAGAGAAAGACGAGAGCGCTTCTCTATATGGTTAGG  479
Db      477  CCATCGCGACATCTACGAGCAGAGAAAGACGAGAGCGCTTCTCTATATGGTTAGG  536
UY      480  CCTCCGAGAAACCTTCGCTTCTGAGCCAGCATAGGSGGCGCTCGGCTGGGAGTGGG  539
Db      537  CCTCCGAGAAACCTTCGCTTCTGAGCCAGCATAGGSGGCGCTCGGCTGGGAGTGGG  596
UY      540  GGGGCGCGGTCAGGCGCTGCCAGAGAGCTTGGTTCGTAAGTGAAGTGGCTTACGG  599
Db      597  GGGGCGCGGTCAGGCGCTGCCAGAGAGCTTGGTTCGTAAGTGAAGTGGCTTACGG  656
UY      600  TGGTGGCTGGGCGAGGCGATGTGCCCCCTAGTCAAGGSGCA  640
Db      657  TGGTGGCTGGGCGAGGCGATGTGCCCCCTAGTCAAGGSGCA  697

RESULT 10
US-60-323-843-7194
: Sequence 7194, Application US/60323843
: GENERAL INFORMATION:
: APPLICANT: Morris, MacDonald
: APPLICANT: Lal, Preeti
: APPLICANT: Diop, Dinh
: TITLE OF INVENTION: METHOD FOR THE IDENTIFICATION OF SEQUENCE POLYMORPHISMS USING
: TITLE OF INVENTION: POLYMORPHISM IDENTIFIED THEREBY
: FILE REFERENCE: GX-0021-1 P
: CURRENT FILING DATE: 2001-09-21
: NUMBER OF SEQ ID NOS: 12060
: SOFTWARE: PERL Program
: SEQ ID NO 7194
: LENGTH: 1009
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 247776.13
US-60-323-843-7194

```

Query Match	98.08;	Score 627.4;	DB 87;	Length 1009;
Best Local Similarity	99.78;	Pred. No. 3.4e-107;		
Matches 639;	Conservative 0;	Mismatches 1;	Indels 1;	Gaps 1;

57 CTCTCCGACGCGCAGCGCGCCGTGTCTCAGCGAGAGCCCGGAGGCCCTTTGAGCGCGAGCGCG 60
 116

61 GAGAG-TCCTGGAGACCTGTAAGACATATCCCGGCTCCGACAGGCTCCGCTCTGG 119
1111
117 GGAGCGCCCGGAGACCCCAACCGGAGATACATATCCCGGCGCCGACAGGCTCCGCTCTGG 176
120 GCGCCAGGCGGAGCCCGCGGATGCGCTTCAGACCGGGGCTTTGAGGCGAGGCGGAGCTTGG 179
177 GCGCCAGGCGGAGCCCGCGGATGCGCTTCAGACCGGGGCTTTGAGGCGAGGCGGAGCTTGG 236
180 GCGACCCCTGTAAAGAGGATACAGACATATCCCGGACGAGCCCGAGTAAATTCCTTGA 239
237 CGAGACCGGTGTAAAGAGGATACAGACATATCCCGGACGAGACCCGAGTAAATTCCTTGA 296
240 TCATCGAGCGGCTTACAAAGGCTGAGAGTACGTGCGGCTCTGGAGCAAGACTAAATTTGG 299
297 TCATCGAGCGGCTTACAAAGGCTGAGAGTACGTGCGGCTCTGGAGCAAGACTAAATTTGG 356
300 TCCCGGACCATGTGCAACATGAGGAGATGGTAAAGATCATGCGGCGCGCGCTGATGATGA 359
357 TCCCGGACCATGTGCAACATGAGGAGATGGTAAAGATCATGCGGCGCGCGCTGATGATGA 416
360 ACCCGACGCGAGGCTCTCTTCCTGCTGGTGAACCCAGACAGCATGAGTGGTATGCTTGA 419
417 ACCCGACGCGAGGCTCTCTTCCTGCTGGTGAACCCAGACAGCATGAGTGGTATGCTTGA 476
420 CCATCGCGGACATCTACGAGCGAGGAGAAAGACGAGACGGCTTCTCTATATGATTAATG 479
477 CCATCGCGGACATCTACGAGCGAGGAGAAAGACGAGACGGCTTCTCTATATGATTAATG 536
480 CCTCCGAGGAAACCTTCGAGCTTCTGAGCGAGTACGTAGGGGGGCTGCGGCTGGTAAATTCGG 539
537 CCTCCGAGGAAACCTTCGAGCTTCTGAGCGAGTACGTAGGGGGGCTGCGGCTGGTAAATTCGG 596
540 GGGGCCCCGGTCAAGCCCTCCCGAGAGAGCTTCTGCTTCGAACTGAGCTGGCTCTATAG 599
597 GGGGCCCCGGTCAAGCCCTCCCGAGAGAGCTTCTGCTTCGAACTGAGCTGGCTCTATAG 656
600 TGGTGGGCTGGGCGAGGAGATATGGCCGCTAGTCAAGAGGCTA 640
657 TGGTGGGCTGGGCGAGGAGATATGGCCGCTAGTCAAGAGGCTA 697

```

? RESULT 11
? US-10-405-027-30
? Sequence 30, Application US/10405027
? GENERAL INFORMATION:
? APPLICANT: Rosen et al.
? TITLE OF INVENTION: Human Secreted Proteins
? FILE REFERENCE: P8806P1
? CURRENT APPLICATION NUMBER: US/10/405,027
? CURRENT FILING DATE: 2003-04-07
? PRIOR APPLICATION NUMBER: 60/459,608
? PRIOR FILING DATE: 2002-04-04
? PRIOR APPLICATION NUMBER: 60/479,175
? PRIOR FILING DATE: 2002-04-30
? NUMBER OF SEQ ID NOS: 5810
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 30
? LENGTH: 1019
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-10-405-027-30

```

	Query Match	98.0%	Score 627.4	DB 52	Length 1017	
	Best Local Similarity	99.7%	Pred. No. 3,4e-107			
	Matches 639; Conservative	0	Mismatches 1	Indels 1	Gaps 1	
QY	1	CTCCGCGACGCCGAGTGGGTCGTATATTTAAGCCCCGGAAGCATTGAGGCATTAATTTTGATTTT				

61 GGAG-GTCCGGAGGCTCCGAAACGG-262N*ATATCCCGCGGGCGGAGAGCGTGTGTTG 119

Db 85 GGAGCCCCGGAGCCCCCAAAACCGACAGACATCCCGCCGCCAGAGCCCCGGCCCTCG 144
 QY 120 CGCCAGGCGGGGCCCCCGCGATGCCCTGAGACCGGCTTTTGAAGAGCGCGAGCTTCG 179
 Db 145 CGCCAGGCGGGGCCCCCGCGATGCCCTGAGACCGGCTTTTGAAGAGCGCGAGCTTCG 204
 QY 180 CGAGCCGCTGTAAAGAGGTAGACAGATCCGGGACAGACCCCGCAAAATCCGGTGA 239
 Db 205 CGAGCCGCTGTAAAGAGGTAGACAGATCCGGGACAGACCCCGCAAAATCCGGTGA 264
 QY 240 TCATGAGACGCTACAAAGGTAGAGAGAGCTGCCGCTCTGAGACAAGACCAAGTTTGG 299
 Db 265 TCATGAGACGCTACAAAGGTAGAGAGAGCTGCCGCTCTGAGACAAGACCAAGTTTGG 324
 QY 300 TCCCGGACCATGTCAACATGAGCGAGTGTCAAGATATCCGGGCGCGCTGACAGTGA 359
 Db 325 TCCCGGACCATGTCAACATGAGCGAGTGTCAAGATATCCGGGCGCGCTGACAGTGA 384
 QY 360 ACCCAGCGAGGCTTTCTTCTGCTGTGAACCAAGACAGCATGGTGTGTCCAGCG 419
 Db 385 ACCCAGCGAGGCTTTCTTCTGCTGTGAACCAAGACAGCATGGTGTGTCCAGCG 444
 QY 420 CCATGCGGGACATCTACAGAGAGAGAGAGAGAGAGCGCTTCTCTATATGTCTACG 479
 Db 445 CCATGCGGGACATCTACAGAGAGAGAGAGAGAGAGCGCTTCTCTATATGTCTACG 504
 QY 480 CCTCCAGGAACCTTGGCTTTCTGAGCGAGAGAGAGAGAGAGAGCGCTTCTCTATATGTCTACG 539
 Db 505 CCTCCAGGAACCTTGGCTTTCTGAGCGAGAGAGAGAGAGAGAGCGCTTCTCTATATGTCTACG 564
 QY 540 GGGCCCCGGTGAAGGCTTCCCGAGAGAGCTTCTGTTCTTGAACCTGAGCTGCTTACCG 599
 Db 565 GGGCCCCGGTGAAGGCTTCCCGAGAGAGAGCTTCTGTTCTTGAACCTGAGCTGCTTACCG 624
 QY 600 TGTGTGGCTGGGAGGAGCATGTGCCCCCTAGTCAGAGGGCA 640
 Db 625 TGTGTGGCTGGGAGGAGCATGTGCCCCCTAGTCAGAGGGCA 665

RESULT 12

US-09-428-151A-9451
 : Sequence 9451, Application US/09428151A
 : GENERAL INFORMATION:
 : APPLICANT: Glucksmann, M. Alexandra
 : TITLE OF INVENTION: Nucleic Acid Molecules Derived from a
 : FILE REFERENCE: 1600.1003001
 : CURRENT APPLICATION NUMBER: US/09/428,151A
 : PRIOR FILING DATE: 1999-10-27
 : PRIOR APPLICATION NUMBER: 60/106,057
 : NUMBER OF SEQ ID NOS: 11036
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 9451
 : LENGTH: 1356
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : US-09-428-151A-9451

Query Match 98.0%; Score 627.4; DB 21; Length 1356;
 Best Local Similarity 99.7%; Pred. No. 3.4e-107;
 Matches 639; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTCCCGAGCGCGAGCGCGCTGCTCAAGCGAGAGCGCGAGCGCTTGAAGCGAGGGCG 60
 Db 391 CTCCCGAGCGCGAGCGCGCTGCTCAAGCGAGAGCGCGAGCGCTTGAAGCGAGGGCG 450
 QY 61 GGAG-CCCGGAGCGCGCGCGCGAGACATCCCGCGCGCGCGCGAGAGCGCGCGCTTCG 119
 Db 451 GGAGCGCGCGCGCGCGCGCGCGCGAGACATCCCGCGCGCGCGCGAGAGCGCGCGCTTCG 510
 QY 120 CGCCAGCGCGCGCGCGCGCGCTGCTCAAGCGAGCGCGCTTGAAGCGAGGGCGAGCTTCG 179

Db 511 CGCCAGCGGGGCCCCCGCGATGCCCTGAGACCGGCTTTTGAAGAGCGCGAGAGCTTCG 570
 QY 180 CGAGCCGCTGTAAAGAGGTAGACAGATCCGGGACAGACCCCGCAAAATCCGGTGA 239
 Db 571 CGAGCCGCTGTAAAGAGGTAGACAGATCCGGGACAGACCCCGCAAAATCCGGTGA 630
 QY 240 TCATGAGACGCTACAAAGGTAGAGAGAGCTGCCGCTCTGAGACAAGACCAAGTTTGG 299
 Db 621 TCATGAGACGCTACAAAGGTAGAGAGAGCTGCCGCTCTGAGACAAGACCAAGTTTGG 690
 QY 300 TCCCGGACCATGTCAACATGAGCGAGTGTCAAGATATCCGGGCGCGCTGACAGTGA 359
 Db 691 TCCCGGACCATGTCAACATGAGCGAGTGTCAAGATATCCGGGCGCGCTGACAGTGA 750
 QY 360 ACCCAGCGAGGCTTTCTTCTGCTGTGAACCAAGACAGCATGGTGTGTCCAGCG 419
 Db 751 ACCCAGCGAGGCTTTCTTCTGCTGTGAACCAAGACAGCATGGTGTGTCCAGCG 810
 QY 420 CCATGCGGGACATCTACAGAGAGAGAGAGAGAGAGCGCTTCTCTATATGTCTACG 479
 Db 811 CCATGCGGGACATCTACAGAGAGAGAGAGAGAGAGCGCTTCTCTATATGTCTACG 870
 QY 480 CCTCCAGGAACCTTGGCTTTCTGAGCGAGAGAGAGAGAGAGAGCGCTTCTCTATATGTCTACG 539
 Db 871 CCTCCAGGAACCTTGGCTTTCTGAGCGAGAGAGAGAGAGAGAGCGCTTCTCTATATGTCTACG 930
 QY 540 GGGCCCCGGTGAAGGCTTCCCGAGAGAGCTTCTGTTCTTGAACCTGAGCTGCTTACCG 599
 Db 931 GGGCCCCGGTGAAGGCTTCCCGAGAGAGAGCTTCTGTTCTTGAACCTGAGCTGCTTACCG 990
 QY 600 TGTGTGGCTGGGAGGAGCATGTGCCCCCTAGTCAGAGGGCA 640
 Db 991 TGTGTGGCTGGGAGGAGCATGTGCCCCCTAGTCAGAGGGCA 1021

RESULT 13

PCT-US98-15949-94
 : Sequence 94, Application PC/TUS9815949
 : GENERAL INFORMATION:
 : APPLICANT: Human Genome Sciences, Inc., et al.
 : TITLE OF INVENTION: P2012PCT
 : FILE REFERENCE: P2012PCT
 : CURRENT APPLICATION NUMBER: PCT/US98/15949
 : PRIOR FILING DATE: 1998-07-29
 : NUMBER OF SEQ ID NOS: 219
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 94
 : LENGTH: 951
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: SITE
 : LOCATION: (826)
 : OTHER INFORMATION: n equals a,t,g, or c
 : PCT-US98-15949-94

Query Match 98.0%; Score 627; DB 1; Length 951;
 Best Local Similarity 99.5%; Pred. No. 4.1e-107;
 Matches 638; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTCCCGAGCGCGAGCGCGCTGCTCAAGCGAGAGCGCGAGCGCTTGAAGCGAGGGCG 60
 Db 6 CTCCCGAGCGCGAGCGCGCTGCTCAAGCGAGAGCGCGAGCGCTTGAAGCGAGGGCG 65
 QY 61 GGAG-CCCGGAGCGCGCGCGCGAGACATCCCGCGCGCGCGCGAGAGCGCGCGCTTCG 119
 Db 66 GGAGCGCGCGCGCGCGCGCGCGAGACATCCCGCGCGCGCGCGAGAGCGCGCGCTTCG 125
 QY 120 CGCCAGCGCGCGCGCGCGCGCTGCTCAAGCGAGCGCGCTTGAAGCGAGGGCGAGCTTCG 179
 Db 126 CGCCAGCGCGCGCGCGCGCGCTGCTCAAGCGAGCGCGCTTGAAGCGAGGGCGAGCTTCG 185
 QY 180 CGAGCCGCTGTAAAGAGGTAGACAGATCCCGAGACAGACCCCGCAAAATCCGGTGA 239

```

1b 186 CCGACCGCTGTAAAGAGGTACAGCAGATCCGGAGCAGACCCCGAGCAAAATCCCGTGA 245
    |||
uy 240 TCATCGACCGCTACAAAGGTGAGAGCAGAGCTCCCGTCTGTGACAGAGTGAAGTTTGG 299
    |||
Db 246 TCATCGACCGCTACAAAGGTGAGAGCAGAGCTCCCGTCTGTGACAGAGTGAAGTTTGG 305
    |||
uy 400 TCCCGGACCATGTCAACATGAGACGAGTTGTCAAGTATCATCGGCGCGCGCTGCAGCTGA 359
    |||
Db 406 TCCCGGACCATGTCAACATGAGACGAGTTGTCAAGTATCATCGGCGCGCGCTGCAGCTGA 365
    |||
uy 460 ACCCGACGAGCGCTTCTTCTGTGTGTGAGACGACAGCATGTGTGATGTCCAGCC 419
    |||
Db 466 ACCCGACGAGCGCTTCTTCTGTGTGAGACGACAGCATGTGTGATGTGTCCAGCC 425
    |||
uy 420 CCATCGCGGACATCTACAGCAGAGGAGAAAGACAGAGCGGTCTCTATATGTGTACG 479
    |||
Db 426 CCATCGCGGACATCTACAGCAGAGGAGAAAGACAGAGCGGTCTCTATATGTGTACG 485
    |||
uy 480 CCTCCAGGAAACCTTCCGCTTCTGTGAGCAGAGTGAAGGCGCTGCGGCTGGAGTCCGG 539
    |||
Db 486 CCTCCAGGAAACCTTCCGCTTCTGTGAGCAGAGTGAAGGCGCTGCGGCTGGAGTCCGG 545
    |||
uy 540 GGGCGCGGCTGAGGCGCTGCGGAGAGGCTTCTGATTCCTGAATGAATGAGCTGACCG 599
    |||
Db 546 GGGCGCGGCTGAGGCGCTGCGGAGAGGCTTCTGATTCCTGAATGAGCTGACCG 605
    |||
uy 600 TGGTGGGCTGGGAGGCAATGCGCGCGCTAGTCAGAGGCA 640
    |||
Db 606 TGGTGGGCTGGGAGGCAATGCGCGCGCTAGTCAGAGGCA 646
    |||

```

RESULT 14

PCT-US98-1549-94

Sequence 94, Application PC/TUS9815949

GENERAL INFORMATION:

APPLICANT: Human Genome Sciences, Inc., et al.

TITLE OF INVENTION: 83 Human Secreted Proteins

FILE REFERENCE: P2012PCT

CURRENT APPLICATION NUMBER: PCT/US98/15949

CURRENT FILING DATE: 1998-07-29

NUMBER OF SEQ ID NOS: 219

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO: 94

LENGTH: 951

TYPE: LNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (826)

OTHER INFORMATION:

PCT-US98-1549-94

Query Match 98.0%; Score 627; DB 2; Length 951;

Best Local Similarity 99.5%; Pred. No. 4.1e-107;

Matches 638; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

```

uy 1 CTCCGGAGCGCGGAGCGCGCTGTCTACAGCGGAGCGCGGAGCGCTTACGCGGAGCGGC 60
    |||
Db 5 CTCCGGAGCGCGGAGCGCGCTGTCTACAGCGGAGCGCGGAGCGCTTACGCGGAGCGGC 65
    |||
uy 61 GGAG-CCCGGAGCGCGGAGCGCGGAGCGGAGCATGCGCGCGCGGAGCGCGGAGCGCTG 119
    |||
Db 66 GGAGCGCGCGGAGCGCGGAGCGCGGAGCGGAGCATGCGCGCGCGGAGCGCGGAGCGCTG 125
    |||
uy 120 GCGCGAGCGCGGCGCGCGGAGTGCCTGAGACCGGCGCTTCAAGTACGCGGAGCGAGCTTG 179
    |||
Db 126 GCGCGAGCGCGGCGCGCGGAGTGCCTGAGACCGGCGCTTCAAGTACGCGGAGCGAGCTTG 185
    |||
uy 180 CCGACCGCTGTAAAGAGGTACAGCAGATCCGGAGCAGACCCCGAGCAAAATCCCGTGA 239
    |||
Db 186 CCGACCGCTGTAAAGAGGTACAGCAGATCCGGAGCAGACCCCGAGCAAAATCCCGTGA 245
    |||

```

```

uy 240 TCATCGACCGCTACAAAGGTGAGAGCAGAGCTCCCGTCTGTGACAGAGTGAAGTTTGG 299
    |||
Db 246 TCATCGACCGCTACAAAGGTGAGAGCAGAGCTCCCGTCTGTGACAGAGTGAAGTTTGG 305
    |||
uy 300 TCCCGGACCATGTCAACATGAGACGAGTTGTCAAGTATCATCGGCGCGCGCTGCAGCTGA 359
    |||
Db 306 TCCCGGACCATGTCAACATGAGACGAGTTGTCAAGTATCATCGGCGCGCGCTGCAGCTGA 365
    |||
uy 360 ACCCGACGAGCGCTTCTTCTGTGTGTGAGACGACAGCATGTGTGATGTCCAGCC 419
    |||
Db 366 ACCCGACGAGCGCTTCTTCTGTGTGTGAGACGACAGCATGTGTGATGTGTCCAGCC 425
    |||
uy 420 CCATCGCGGACATCTACAGCAGAGGAGAAAGACAGAGCGGTCTCTATATGTGTACG 479
    |||
Db 426 CCATCGCGGACATCTACAGCAGAGGAGAAAGACAGAGCGGTCTCTATATGTGTACG 485
    |||
uy 480 CCTCCAGGAAACCTTCCGCTTCTGTGAGCAGAGTGAAGGCGCTGCGGCTGGAGTCCGG 539
    |||
Db 486 CCTCCAGGAAACCTTCCGCTTCTGTGAGCAGAGTGAAGGCGCTGCGGCTGGAGTCCGG 545
    |||
uy 540 GGGCGCGGCTGAGGCGCTGCGGAGAGGCTTCTGATTCCTGAATGAATGAGCTGACCG 599
    |||
Db 546 GGGCGCGGCTGAGGCGCTGCGGAGAGGCTTCTGATTCCTGAATGAGCTGACCG 605
    |||
uy 600 TGGTGGGCTGGGAGGCAATGCGCGCGCTAGTCAGAGGCA 640
    |||
Db 606 TGGTGGGCTGGGAGGCAATGCGCGCGCTAGTCAGAGGCA 646
    |||

```

RESULT 15

US-60-172-373-10400

Sequence 10400, Application US/60172474

GENERAL INFORMATION:

APPLICANT: Morris, MacDonald

APPLICANT: Lal, Preeti

TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using

FILE REFERENCE: GX-000619

CURRENT APPLICATION NUMBER: US/60/172,474

CURRENT FILING DATE: 1994-12-16

NUMBER OF SEQ ID NOS: 25,772

SOFTWARE: PERL Program

SEQ ID NO: 10400

LENGTH: 1017

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc-feature

OTHER INFORMATION: Incycle ID No: 247776.3

FEATURE:

NAME/KEY: unsure

LOCATION: 14, 25, 46

OTHER INFORMATION:

US-60-172-373-10400

Query Match 94.6%; Score 605.4; DB 71; Length 1017;

Best Local Similarity 99.4%; Pred. No. 4.3e-103;

Matches 639; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

```

uy 1 CTCCGGAGCGCGGAGCGCGCTGTCTACAGCGGAGCGCGGAGCGCTTACGCGGAGCGGC 60
    |||
Db 71 CTCCGGAGCGCGGAGCGCGCTGTCTACAGCGGAGCGCGGAGCGCTTACGCGGAGCGGC 140
    |||
uy 61 GGAG-CCCGGAGCGCGGAGCGCGGAGCGGAGCATGCGCGCGCGGAGCGCGGAGCGCTG 117
    |||
Db 131 GGAGCGCGCGGAGCGCGGAGCGCGGAGCGGAGCATGCGCGCGCGGAGCGCGGAGCGCTG 190
    |||
uy 118 GCGCGAGCGCGGCGCGCGGAGTGCCTGAGACCGGCGCTTCAAGTACGCGGAGCGAGCTTG 177
    |||
Db 191 GCGCGAGCGCGGCGCGCGGAGTGCCTGAGACCGGCGCTTCAAGTACGCGGAGCGAGCTTG 250
    |||
uy 178 GCGCGAGCGCGCTGTAAAGAGGTACAGCAGATCCGGAGCAGACCCCGAGCAAAATCCCGT 247
    |||

```

```

Db      251  |||||CGCCACCGCTGTAGAGGTACAGCAGATCCGACACGACCCAGCAAAATCCCGGT 310
QY      238  GATCATCGAGCGCTCTCAAGGGTGAAGAGAGCTGCCCTCCGTGACAAAGCAACTTTT 297
Db      311  GATCATCGAGCGCTCAAGGGTGAAGAGAGCTGCCCTCCGTGACAAAGCAAGTTT 370
QY      298  GGTCCCGGACCATGTCAACATGAGCGAGTGGTCAAGATCATCCGGCCCGCTGCAGCT 357
Db      371  GGTCCCGGACCATGTCAACATGAGCGAGTGGTCAAGATCATCCGGCCCGCTGCAGCT 430
QY      358  GAACCCCAAGCAGCGCTTCTCTCTGCTGTAACCAAGCAGCAGATGTGAGTGTCCAC 417
Db      431  GAACCCCAAGCAGCGCTTCTCTCTGCTGTAACCAAGCAGCAGATGTGAGTGTCCAC 490
QY      418  GCCCATCGGACATCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 477
Db      491  GCCCATCGGACATCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 550
QY      478  CGCCTCCAGGAAACCTTCGCTCTGAGCCAGCAGTAGAGGGGCTCGGCTGGAGTGG 537
Db      551  CGCCTCCAGGAAACCTTCGCTCTGAGCCAGCAGTAGAGGGGCTCGGCTGGAGTGG 610
QY      538  GGGGGCCCCGTCAGGGCCCCGTCAGAGAGCTTGTGTTCTGAACTGAGCTGCCTTAC 597
Db      611  GGGGGCCCCGTCAGGGCCCCGTCAGAGAGCTTGTGTTCTGAACTGAGCTGCCTTAC 670
QY      598  CGTGTGGGCTGGGAGGAGATGTGCCCTAGTCAGAGGGCA 640
Db      671  CGTGTGGGCTGGGAGGAGATGTGCCCTAGTCAGAGGGCA 713

```

Search completed: August 15, 2003, 10:21:12
 Job time : 2938 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

MM protein - protein search, using sw model

Run on: August 11, 2003, 09:49:08 ; Search time 81 Seconds
(without alignments)
231.397 Million cell updates/sec

Title: US-09-904-603-1

Perfect score: 625
Sequence: 1 MPSDRPFKQRRSFADRCKEY.....EKDEGFLYVVASQETFGF 121

Scoring table: BLUSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum PH seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcadata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcadata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcadata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSL/gcadata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSL/gcadata/geneseq/geneseq-emb1/AA2003.DAT:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	625	100.0	121	19	AA669979
2	625	100.0	159	23	ABP41404
3	625	100.0	199	22	ABG09793
4	625	83.4	125	22	AA639660
5	625	83.4	125	22	AA674775
6	625	83.4	125	23	ABG79539
7	625	83.4	130	23	ABG97779
8	625	83.4	142	22	AA674776
9	625	83.4	155	22	AA641446
					Human microtubule
					Human ovarian anti
					Novel human diamo
					Human polypeptide
					Human MPL3 protein
					Human tumour suppl
					Human secretory po
					Rat Mpl3 protein.
					Human polypeptide

10	521	83.4	171	21	AA646006	Human: "CAUTER ASSOC
11	466.5	74.5	99	22	ABG09794	Novel human diamo
12	411	65.8	222	22	AA606002	Novel human secret
13	359	57.4	147	21	AA619555	Human cytoskeleton
14	301	48.2	120	22	AA660179	Human protein SBY
15	272	43.5	54	22	ABG56031	Human liver peptid
16	272	43.5	54	22	ABG40581	Peptide #8087 conc
17	272	43.5	54	22	AA661469	Human brain expres
18	272	43.5	54	22	AA674229	Human bone marrow
19	272	43.5	54	22	AA641441	Peptide #8078 conc
20	272	43.5	135	22	AA679195	Human protein SBY
21	263	42.1	66	23	ABG42427	Human secretory pro
22	236	37.8	116	21	AA618966	Zea mays protein 1
23	234	37.4	121	21	AA652262	Arabidopsis thalia
24	233	37.3	151	22	ABG11284	Novel human diamo
25	230	36.8	120	21	AA637418	Arabidopsis thalia
26	228	36.5	119	21	AA644817	Zea mays protein 1
27	227	36.3	109	21	AA618967	Zea mays protein 1
28	227	36.3	120	21	AA642418	Arabidopsis thalia
29	227	36.3	120	21	AA644107	Zea mays protein 1
30	227	36.3	120	21	AA644172	Zea mays protein 1
31	227	36.3	120	21	AA643761	Arabidopsis thalia
32	227	36.3	140	21	AA642416	Arabidopsis thalia
33	226	36.2	122	21	AA642625	Arabidopsis thalia
34	225	36.0	122	21	AA642625	Arabidopsis thalia
35	223	35.7	115	21	AA649168	Arabidopsis thalia
36	223	35.7	115	21	AA649473	Arabidopsis thalia
37	220	35.2	114	21	AA649169	Arabidopsis thalia
38	219	35.0	107	21	AA642625	Arabidopsis thalia
39	219	35.0	122	21	AA608604	Arabidopsis thalia
40	219	35.0	122	21	AA641655	Arabidopsis thalia
41	219	35.0	122	21	AA647988	Arabidopsis thalia
42	215	35.0	122	21	AA647988	Arabidopsis thalia
43	215	35.0	184	21	AA647987	Arabidopsis thalia
44	217	34.7	119	21	AA610031	Arabidopsis thalia
45	217	34.7	119	21	AA611019	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AA669979 standard; Protein: 121 AA.
ID AA669979
AC AA669979:
XX 20-NDV-1998 (first entry)
XX Human microtubule associated protein hM3.
DE Human microtubule associated protein hM3.
XX Human: microtubule-associated protein; hMAP; cell proliferation; cancer.
OS Homo sapiens.
XX
PN WO9837197-A1.
XX
XX 27-AUG-1998.
XX
XX 23-FEB-1998; 98W00980744.
XX
XX 24-FEB-1997; 97US0005117.
XX
XX (INCYTE) INCYTE PHARM INC.
XX Coli SK, Hillman JL:
XX WPI: 1598-467560/40.
XX N-PSDB: AA643780.
XX
XX New isolated human microtubule-associated protein - used to develop
XX products for use in the diagnosis, prevention and treatment of
XX disorders associated with cell proliferation, particularly cancers

XX Claim 1; Fig 1; 58pp; English.
PS
XX
CC Human microtubule-associated protein (hMAP) expression is associated with
CC disorders of cell proliferation, in particular, cancer. The hMAP and
CC agonists can be used to stimulate cell proliferation. Tissues or
CC organs generated by these methods may be introduced or transplanted into
CC a subject for the purpose of treating a variety of diseases. Antagonists
CC or inhibitors of hMAP can be used to treat or prevent disorders
CC associated with cell proliferation, particularly cancers. The products
CC can also be used for detection, diagnosis and drug screening.
XX
SQ Sequence 121 AA;
Query Match 100.0%; Score 625; DB 19; Length 121;
Best Local Similarity 100.0%; Pred. No. 4.9e-65;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPSDRPFKORRSFADRCKEVOQIRDQHPKRPVIERKYGKQOLPVDKTRFLVDPHVM 60
DB 1 MPSDRPFKORRSFADRCKEVOQIRDQHPKRPVIERKYGKQOLPVDKTRFLVDPHVM 60
QY 61 SELVKIIRRLQNLPTQAFLLVNHQSMWSVSPDIADIEQEKDEDFLYWVASQETFG 120
DB 61 SELVKIIRRLQNLPTQAFLLVNHQSMWSVSPDIADIEQEKDEDFLYWVASQETFG 120
QY 121 F 121
DB 121 F 121
Db 121 F 121
RESULT 2
ABP41404
ID ABP41404 standard; Protein; 169 AA.
XX
AC ABP41404;
XX
DT 22-AUG-2002 (first entry)
XX
DE Human ovarian antigen HFATL31, SEQ ID NO:2536.
XX
KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive.
XX
OS Homo sapiens.
XX
PN WO20020677-A1.
XX
PD 03-JAN-2002.
XX
PF 07-JUN-2001; 2001WO-US18569.
XX
PR 07-JUN-2000; 2000US-209467P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-147878/19.
XX
DR N-PSDB; ABQ54481.
XX
PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -

XX Claim 11; SEQ ID NO 2536; 2922pp; English.
PS
XX
CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosting or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents a human ovarian antigen of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 169 AA;
Query Match 100.0%; Score 625; DB 23; Length 169;
Best Local Similarity 100.0%; Pred. No. 7.6e-65;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPSDRPFKORRSFADRCKEVOQIRDQHPKRPVIERKYGKQOLPVDKTRFLVDPHVM 60
DB 49 MPSDRPFKORRSFADRCKEVOQIRDQHPKRPVIERKYGKQOLPVDKTRFLVDPHVM 108
QY 61 SELVKIIRRLQNLPTQAFLLVNHQSMWSVSPDIADIEQEKDEDFLYWVASQETFG 120
DB 109 SELVKIIRRLQNLPTQAFLLVNHQSMWSVSPDIADIEQEKDEDFLYWVASQETFG 168
QY 121 F 121
DB 169 F 169
RESULT 3
ABG09793
ID ABG09793 standard; Protein; 199 AA.
XX
AC ABG09793;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #9784.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PR 30-MAR-2001; 2001WO-US08631.
XX

PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSF-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB: AAS73980.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20: SEQ ID NO 40152; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (II) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AB000010-AB030377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 199 AA;
Query Match 100.0%; Score 625; DB 22; Length 199;
Best Local Similarity 100.0%; Pred. No. 9.4e-65;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Caps 0;
QY 1 MSPDRPKORRSFADCKEYQOIRDOHPSKIPVIERKKEKOLPVDKTKFLVPPHVM 60
Db 52 MSPDRPKORRSFADCKEYQOIRDOHPSKIPVIERKKEKOLPVDKTKFLVPPHVM 111
QY 61 SELVKIIRRLQINPTQAFLLVNGHSMVSVSTPIADIYQEQDEDEGFLMNVASQETFG 120
Db 112 SELVKIIRRLQINPTQAFLLVNGHSMVSVSTPIADIYQEQDEDEGFLMNVASQETFG 171
QY 121 F 121
Db 172 F 172
RESULT 4
AAM39660
ID AAM39660 standard; Protein: 125 AA.
XX
AC AAM39660;
XX
DT 22-OCT-2001 (first entry)
DE Human polypeptide SEQ ID NO 2805.
XX
KW Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;

KW Leukaemia.
XX
OS Homo sapiens.
XX
PN WO20015312-A1.
XX
PD 26-JUL-2001.
XX
PE 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSF-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang Y;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR N-PSDB: AAI58816.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 4: SEQ ID NO 2805; 1007pp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161365) and
CC the encoded polypeptides (AAM38642-AAM42213) with noctropic.
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 125 AA;
Query Match 83.4%; Score 521; DB 22; Length 125;
Best Local Similarity 82.5%; Pred. No. 7.2e-53;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Caps 0;
QY 1 MSPDRPKORRSFADCKEYQOIRDOHPSKIPVIERKKEKOLPVDKTKFLVPPHVM 60
Db 1 MSPDKTFKORRTQOREVEDRLIRDOHPIKIPVIERKKEKOLPVDKTKFLVPPHVM 60
QY 61 SELVKIIRRLQINPTQAFLLVNGHSMVSVSTPIADIYQEQDEDEGFLMNVASQETFG 120
Db 61 SELVKIIRRLQINPTQAFLLVNGHSMVSVSTPIADISYSEKDEDEGFLMNVASQETFG 120
RESULT 5
AAB74775
ID AAB74775 standard; Protein: 125 AA.
XX
AC AAB74775;
XX
DT 06-JUN-2001 (first entry)
DE

DE Human MPL3 protein SEQ ID NO:4.
XX
XX Human; microtubule-associated protein 1A/1B light chain 3; MPL3;
KW hMPL3.
XX
XX Homo sapiens.
OS
XX CNI274726-A.
PN
XX 29-NOV-2000.
PD
XX
XX 24-MAY-1999; 99CN-0106971.
PE
XX
XX 24-MAY-1999; 99CN-0106971.
PR
XX
XX (UYFU-) UNIV FUDAN.
PA
XX
XX Yu L, Zhao Y, Fu Q;
PI
XX
XX WPI; 2001-211747/22.
DR
XX N-PSDB; AAF81833.
DR
XX
XX Human tubule related protein 1A/1B light chain 3 and its code sequence,
PT preparation and application -
XX
XX Claim 4: Page 15; 20pp; Chinese.
PS
XX
XX The present invention describes the human microtubule-associated protein
CC 1A/1B light chain 3 (hMPL3), which is a homologue of rat MPL3. The
CC present invention also describes a method for the application and
CC production process of the hMPL3 polynucleotide and protein sequences.
CC The present sequence represents the hMPL3 protein as given in the
CC present invention.
XX
XX Sequence 125 AA;
SQ

Query Match 83.4%; Score 521; DB 22; Length 125;
Best Local Similarity 82.5%; Pred. No. 7.2e-53;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

OY 1 MDSRPFKORSEFADCKEVOQIRDOHPKIRPIIIRYKGEKQLPVLDTKFLVDPDHVM 60
DB 1 MPSEKTFKORRFEQREVDRLIRQHPRTKIPVIERKYGKQLPVLDTKFLVDPDHVM 60
OY 61 SELVVIIRRLQNLNTQAFELLYNOSHWSVSTPIADIYEOBKDEGFLYVYASQETFG 120
DB 61 SELIKIIRRLQNLNQAFLLYNOSHWSVSTPISEVYESKDEGFLYVYASQETFG 120

RESULT 6
ABG79539 standard; Protein; 125 AA.
ID ABG79539;
AC
XX
XX ABG79539;
DT
XX 15-NOV-2002 (first entry)
DE
XX Human tumour suppressor protein #25.
XX
XX Human; cancer; tumour suppressor protein; breast cancer; lymphoma;
KW adenocarcinoma; leukaemia; melanoma; myeloma; sarcoma; teratocarcinoma;
KW prostate; cervix; liver; ovary; adrenal gland; heart; brain; lung; colon;
KW placenta; skeletal muscle; synovial membrane; tonsil; kidney; uterus;
KW skin; cytotlastic.
XX
XX Homo sapiens.
OS
XX WO300264775-A1.
PN
XX 22-AUG-2002.
PD
XX
XX 12-FEB-2002; 2002WO-AU00137.
PE
XX

PR 12-FEB-2001; 2001AU-0003052.
PR 12-FEB-2001; 2001AU-0003053.
PR 12-FEB-2001; 2001AU-0003054.
XX
XX (BION-) BIONOMICS LTD.
PA
XX
XX Callen DF, Powell JA, Kremmidiotis G, Gardner AE, Whitmore SA;
PI
XX
XX WPI; 2002-657597/70.
DR
XX N-PSDB; ABS64852.
DR
XX
XX New tumour suppressor genes identified at 16q24.3, useful for
PT identifying or obtaining full-length human genes involved in the
PT tumorigenic process, or in diagnosing or treating cancer (e.g. breast
PT or cervix cancer) through gene therapy -
XX
XX Claim 22: Page 205; 218pp; English.
PS
XX
XX The invention relates to a tumour suppressor gene and the protein it
CC encodes. The genes are useful for identifying and/or obtaining
CC full-length human genes involved in the tumorigenic process. The genes
CC the polypeptides and antibodies to the polypeptides are useful in
CC diagnosing cancer, in establishing the prognosis of a patient diagnosed
CC with cancer, or in treating cancer through gene therapy. In particular,
CC the cancer is breast cancer. The drugs that restore tumour suppressor
CC activity are useful for treating breast cancer or for manufacturing a
CC medicament for the treatment of breast cancer. The cancer may also be
CC adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma, sarcoma,
CC teratocarcinoma, or cancer of the prostate, cervix, liver, ovary, adrenal
CC gland, heart, brain, lung, placenta, skeletal muscle, synovial membrane,
CC tonsil, kidney, colon, uterus or skin. This sequence represents a human
CC tumour suppressor protein of the invention.
XX
XX Sequence 125 AA;
SQ

Query Match 83.4%; Score 521; DB 23; Length 125;
Best Local Similarity 82.5%; Pred. No. 7.2e-53;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

OY 1 MDSRPFKORSEFADCKEVOQIRDOHPKIRPIIIRYKGEKQLPVLDTKFLVDPDHVM 60
DB 1 MPSEKTFKORRFEQREVDRLIRQHPRTKIPVIERKYGKQLPVLDTKFLVDPDHVM 60
OY 61 SELVVIIRRLQNLNTQAFELLYNOSHWSVSTPIADIYEOBKDEGFLYVYASQETFG 120
DB 61 SELIKIIRRLQNLNQAFLLYNOSHWSVSTPISEVYESKDEGFLYVYASQETFG 120

RESULT 7
ABB97779 standard; Protein; 130 AA.
ID ABB97779;
AC
XX
XX ABB97779;
DT
XX 03-OCT-2002 (first entry)
DE
XX Human secretory polypeptide (SPTM) 31.
XX
XX Human; secretory protein; secretory polynucleotides; SPTM;
KW SPTM-related disease; somatic gene therapy; germline gene therapy;
KW severe combined immunodeficiency; intracellular parasite protection;
KW fungal parasite; protozoan parasite; cell proliferative disorder; cancer;
KW immune disorder; AIDS; neurological disorder; multiple sclerosis;
KW motor neuron disorder; demyelinating disease; Parkinson's disease;
KW meningitis; abscess; prion diseases; cerebral palsy;
KW neuroskeletal disorder; peripheral nervous system disorder;
KW dermatomyositis; polymyositis; myopathy; myasthenia gravis;
KW mental disorder; Tourette's syndrome.
XX
XX Homo sapiens.
OS
XX WO200220756-A2.
PN
XX

PD 14-MAR-2002.
 XX
 EE 30-AUG-2001: 2001WO-US27297.
 XX
 PR 05-SEP-2000: 2000US-229747P.
 PR 05-SEP-2000: 2000US-229748P.
 PR 05-SEP-2000: 2000US-229749P.
 PR 05-SEP-2000: 2000US-229750P.
 PR 05-SEP-2000: 2000US-229751P.
 PR 05-SEP-2000: 2000US-230016P.
 PR 05-SEP-2000: 2000US-230583P.
 PR 06-SEP-2000: 2000US-230505P.
 PR 06-SEP-2000: 2000US-230514P.
 PR 06-SEP-2000: 2000US-230515P.
 PR 06-SEP-2000: 2000US-230517P.
 PR 06-SEP-2000: 2000US-230518P.
 PR 06-SEP-2000: 2000US-230519P.
 PR 06-SEP-2000: 2000US-230595P.
 PR 06-SEP-2000: 2000US-230596P.
 PR 06-SEP-2000: 2000US-230597P.
 PR 06-SEP-2000: 2000US-230599P.
 PR 06-SEP-2000: 2000US-230610P.
 PR 06-SEP-2000: 2000US-230864P.
 PR 06-SEP-2000: 2000US-230865P.
 PR 06-SEP-2000: 2000US-230988P.
 PR 06-SEP-2000: 2000US-230989P.
 PR 07-SEP-2000: 2000US-230950P.
 PR 07-SEP-2000: 2000US-230897P.
 PR 07-SEP-2000: 2000US-230951P.
 PR 07-SEP-2000: 2000US-231163P.
 PR 07-SEP-2000: 2000US-231832P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;
 PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;
 PI Momiyama EG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;
 PI Gerstin MG, Peralta CH, David MH, Panzer SR, Flores V, Datto A;
 PI Marwaha R, Chen AJ, Chang SC, Au AP, Iman RR;
 XX
 DR WPI: 2002-315658/35.
 DR N-PSDB: ABL99776.
 XX
 PT Polynucleotide sequences encoding human secretory proteins useful for
 PT gene therapy of e.g. genetic deficiency disorders, cancers, and
 PT diseases caused by intracellular parasites -
 XX
 PS Claim 29: Page 385: 585pp: English.
 XX
 CC The invention comprises the amino acid and coding sequences of human
 CC secretory (SPTM) proteins. The SPTM DNA and amino acid sequences are
 CC useful for treating a disease or condition associated with the expression
 CC of functional SPTM. The SPTM DNA sequences are useful for somatic or
 CC germ-line gene therapy to correct a genetic deficiency (e.g. severe
 CC combined immunodeficiency). The SPTM DNA sequences are also useful in
 CC providing protection against intracellular parasites (e.g. fungal
 CC parasites and protozoan parasites). The SPTM DNA and protein sequences
 CC are also useful for diagnosing cell proliferative disorders, cancer,
 CC immune disorders (e.g. AIDS), neurological disorders (e.g. Parkinson's
 CC disease), motor neuron disorders, demyelinating diseases (e.g. multiple
 CC sclerosis), meningitis, abscesses, prion diseases, cerebral palsy,
 CC neuroskeletal disorders, peripheral nervous system disorders,
 CC dermatomycosis and polymyositis, myopathy, myasthenia gravis, and mental
 CC disorders (e.g. Tourette's syndrome). Amino acid sequences ABB97749 -
 CC ABB97943 represent human secretory proteins of the invention.
 XX
 SQ Sequence 130 AA:
 Query Match 83.4%: Score 521: DH 23: Length 130:
 Best Local Similarity 82.5%: Pred. No. 7.6e-53:
 Matches 99: Conservative 12: Mismatches 9: Indels 0: Gaps 0:

QY 1 MSPDRPEKRRSFADCKEYVOI RDOHPSKIPVILIERKSGEKQLPVLDKIKELVDPHYNM 60
 DB 6 MPSEKTFKQKRITFEEDKRVAVLKEGHHTKIPVILIERKSGEKQLPVLDKIKELVDPHYNM 65
 QY 61 SELVKIIRRKQLQLNFGAFILLVNQHSMVSVSTPLADIVEDQKDEGFLVWVVASQELP 120
 DB 66 SELIKIIRRKQLNANQAFILLVNQHSMVSVSTPISEYSESEKDEGFLVWVVASQELP 125
 RESULT 8
 AAB74776
 ID AAB74776 standard: protein: 142 AA.
 XX
 AC AAB74776:
 XX
 DT 06-JUN-2001 (first entry)
 XX
 DE Rat MPL3 protein.
 XX
 KW Human: microtubule-associated protein 1A/1B light chain 3; MPL3;
 KW hMPL3.
 OS Rattus sp.
 XX
 PN CN1274726-A.
 PD 29-NOV-2000.
 XX
 PF 24-MAY-1999: 99CN-0106971.
 XX
 PR 24-MAY-1999: 99CN-0106971.
 XX
 PA (UYFU-) UNIV FUJIAN.
 XX
 PI Yu L, Zhao Y, Fu Y;
 DR WPI: 2001-211747/22.
 XX
 PT Human tubule related protein 1A/1B light chain 3 and its cDNA sequence,
 PT preparation and application -
 XX
 PS Example 2: Fig 1: 20pp: Chinese.
 XX
 CC The present invention describes the human microtubule-associated protein
 CC 1A/1B light chain 3 (hMPL3), which is a homologue of rat MPL3. The
 CC present invention also describes a method for the application and
 CC production process of the hMPL3 polynucleotide and protein sequences.
 CC The present sequence represents the rat MPL3 protein which is given
 CC in comparison with the hMPL3 protein in an example from the present
 CC invention.
 XX
 SQ Sequence 142 AA:
 Query Match 83.4%: Score 521: DH 22: Length 142:
 Best Local Similarity 82.5%: Pred. No. 8.5e-53:
 Matches 99: Conservative 12: Mismatches 9: Indels 0: Gaps 0:
 QY 1 MPSIRPKRQKRSFADCKEYVOI RDOHPSKIPVILIERKSGEKQLPVLDKIKELVDPHYNM 60
 DB 1 MPSEKTFKQKRITFEEDKRVAVLKEGHHTKIPVILIERKSGEKQLPVLDKIKELVDPHYNM 65
 QY 61 SELVKIIRRKQLQLNFGAFILLVNQHSMVSVSTPLADIVEDQKDEGFLVWVVASQELP 120
 DB 61 SELIKIIRRKQLNANQAFILLVNQHSMVSVSTPISEYSESEKDEGFLVWVVASQELP 120
 RESULT 9
 AAM41446
 ID AAM41446 standard: protein: 155 AA.
 XX
 AC AAM41446:
 XX
 DT 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 6377.
DE
XX
XX Human: nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
XX Homo sapiens.
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0683036.
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSEQ -) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang JT, Zhou Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou F, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
XX N-PSDB: AAI60602.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 6377; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 155 AA;
SQ
Query Match 83.4%; Score 521; DB 22; Length 155;
Best Local Similarity 82.5%; Pred. No. 9.6e-53;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;
OY 1 MSPDRPKRRSFADCKEYQVQIRDOHPSKIPVIIEYKKGKOLPVLDKTKFLVDPHVM 60
DB 31 MPSEKTKORRTEQORVEDVRLREQHPRIPIVIEYKKGKOLPVLDKTKFLVDPHVM 90
OY 61 SELVKKIRRLQINPTQAFLLVNOHSMVSVPIDIVYEOEKDEDFLVMVYASQETFG 120
DB 91 SELIKIRRLQINANQAFLLVNGHSMVSVPISVYSEKDEDFLVMVYASQETFG 150

RESULT 10
AAB43606
ID AAB43606 standard; Protein: 171 AA.
XX
XX AAB43606;
AC
XX 08-FEB-2001 (first entry)
DT
XX
XX Human cancer associated protein sequence SEQ ID NO:1051.
DE
XX
XX Human: cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
KW antidiabetic; antiaesthetic; antirheumatic; antiarthritic; antiviral;
KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening.
XX
XX Homo sapiens.
XX WO200053350-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US05882.
XX
XX 12-MAR-1999; 99US-0124270.
XX
XX (HUMA -) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
PI
XX
XX WPI: 2000-587533/55.
XX N-PSDB: AAC77815.
XX
XX Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer -
XX
XX Claim 11; Page 1642-1643; 2352pp; English.
XX
XX AAC77607 to AAC78448 encode the human cancer associated proteins given
CC in AAB43398 to AAB44233. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnery; immunomodulator;
CC antidiabetic; antiaesthetic; antirheumatic; antibacterial; antiviral;
CC antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat. Immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention.
XX
XX Sequence 171 AA;
SQ
Query Match 83.4%; Score 521; DB 21; Length 171;
Best Local Similarity 82.5%; Pred. No. 1.1e-52;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;
OY 1 MSPDRPKRRSFADCKEYQVQIRDOHPSKIPVIIEYKKGKOLPVLDKTKFLVDPHVM 60

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 10:12:08 : Search time 30 Seconds
(without alignments)
170.654 Million cell updates/sec

Title: US-09-904-603-1
Perfect score: 625
Sequence: 1 MPSDRPKQRKRSFADCKEY.....EKDEDGLVMYASQETFG 121

Scoring table:
BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 4210858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:★
1: /cgn2-6/plodata/1/1aa/5A-COMB.pep:★
2: /cgn2-6/plodata/1/1aa/5B-COMB.pep:★
3: /cgn2-6/plodata/1/1aa/6A-COMB.pep:★
4: /cgn2-6/plodata/1/1aa/6B-COMB.pep:★
5: /cgn2-6/plodata/1/1aa/PCTUS-COMB.pep:★
6: /cgn2-6/plodata/1/1aa/backfilltest.pep:★

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	625	100.0	121	2	US-08-805-117-1 Sequence 1, Appli
2	625	100.0	121	3	US-09-199-838-1 Sequence 1, Appli
3	521	83.4	142	2	US-08-805-117-3 Sequence 3, Appli
4	521	83.4	142	3	US-09-199-838-3 Sequence 4, Appli
5	267	42.7	68	4	US-09-708-725A-4 Sequence 2, Appli
6	263	42.1	66	4	US-09-708-725A-2 Sequence 9, Appli
7	77.5	12.4	724	4	US-08-671-757A-9 Sequence 25, Appli
8	70	11.2	527	3	US-08-369-822C-25 Sequence 40, Appli
9	70	11.2	527	3	US-08-582-776C-40 Sequence 37, Appli
10	70	11.2	589	2	US-08-434-831B-37 Sequence 6, Appli
11	70	11.2	590	4	US-09-398-395A-54 Sequence 54, Appli
12	70	11.2	590	4	US-09-887-586A-54 Sequence 54, Appli
13	70	11.2	590	4	US-09-895-752-54 Sequence 54, Appli
14	70	11.2	590	4	US-09-903-012B-54 Sequence 54, Appli
15	69	11.0	704	1	US-08-188-582-5 Sequence 62, Appli
16	69	11.0	704	1	US-08-188-582-5 Sequence 5, Appli
17	69	11.0	704	1	US-08-646-715-5 Sequence 3, Appli
18	69	11.0	704	2	US-08-308-818-3 Sequence 62, Appli
19	69	11.0	704	3	US-08-477-346-62 Sequence 62, Appli
20	69	11.0	704	4	US-08-473-089-62 Sequence 62, Appli
21	69	11.0	704	4	US-08-487-072A-62 Sequence 62, Appli
22	68	10.9	251	4	US-09-107-532A-5498 Sequence 9, Appli
23	68	10.9	251	4	US-08-335-865J-9 Sequence 10, Appli
24	68	10.9	251	4	US-08-913-159-10 Sequence 5455, Ap
25	67.5	10.8	319	4	US-09-134-001C-5455 Sequence 5082, Ap
26	67	10.7	319	4	US-09-328-352-5082
27	66.5	10.6	387	4	US-09-328-352-5082

28	66.5	10.6	455	4	US-09-252-991A-18265 Sequence 14, Appli
29	66.5	10.6	626	3	US-09-357-251-34 Sequence 75, Appli
30	66.5	10.6	648	4	US-09-370-861A-75 Sequence 2497, Ap
31	66	10.6	564	4	US-09-134-001C-2897 Sequence 12, Appli
32	65.5	10.5	322	4	US-09-359-161-7 Sequence 12, Appli
33	65.5	10.5	4066	3	US-08-952-127-12 Sequence 6668, Ap
34	65	10.4	486	3	US-08-348-518C-2 Sequence 2, Appli
35	65	10.4	872	4	US-09-107-532A-6668 Sequence 2, Appli
36	65	10.4	1708	1	US-08-493-092-2 Sequence 8, Appli
37	65	10.4	1708	1	US-08-508-836A-2 Sequence 8, Appli
38	65	10.4	3056	1	US-08-508-836A-8 Sequence 4, Appli
39	65	10.4	3056	2	US-08-629-001A-3 Sequence 4, Appli
40	65	10.4	3056	2	US-08-874-266-2 Sequence 4, Appli
41	65	10.4	3056	3	US-08-642-274D-3 Sequence 4, Appli
42	65	10.4	3056	3	US-08-952-127-3 Sequence 4, Appli
43	65	10.4	3056	3	US-08-952-127-3 Sequence 4, Appli
44	65	10.4	3056	4	US-09-360-416-2 Sequence 4, Appli
45	65	10.4	3057	4	US-09-360-416-3 Sequence 4, Appli

ALIGNMENTS

```
RESULT 1
US-08-805-117-1
: Sequence 1, Application US/08805117
: Patent No. 5955312
:
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Goll, Surya K.
: TITLE OF INVENTION: NOVEL MICROBULE-ASSOCIATED PROTEIN
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Inyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/805,117
: FILING DATE: Bretwlich
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: EP-0211 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
: INFORMATION FOR SHU ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 121 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: THYKNOT14
: CLONE: 144378
: US-08-805-117-1
:
: Query Match 100.0% Score 625, DB 2, Length 121:
: Best local similarity 100.0% Pred. No. 2.2e+69:
: Matches 121: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
```

OY 1 MPSDRPFKORRSFADRCKEVOQIRDOHPSKIPVIERKGEKOLPVLDTKFLVDPDHVM 60
DB 1 MPSDRPFKORRSFADRCKEVOQIRDOHPSKIPVIERKGEKOLPVLDTKFLVDPDHVM 60
OY 61 SELVKIIRRLQLNPTQAFLLVNOHSMVSVTPDIADIEQEKDEGFLVMVYASQETFG 120
DB 61 SELVKIIRRLQLNPTQAFLLVNOHSMVSVTPDIADIEQEKDEGFLVMVYASQETFG 120
OY 121 F 121
DB 121 F 121

RESULT 2
US-09-199-838-1
Sequence 1, Application US/09199838

Patent No. 6280733
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL MICROTUBULE-ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/199,838
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/805,117
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0211 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 121 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: THYRMOT03
CLONE: 144378
US-09-199-838-1

Query Match 100.0%; Score 625; DB 3; Length 121;
Best Local Similarity 100.0%; Pred. No. 2.2e-69;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MPSDRPFKORRSFADRCKEVOQIRDOHPSKIPVIERKGEKOLPVLDTKFLVDPDHVM 60
DB 1 MPSDRPFKORRSFADRCKEVOQIRDOHPSKIPVIERKGEKOLPVLDTKFLVDPDHVM 60
OY 61 SELVKIIRRLQLNPTQAFLLVNOHSMVSVTPDIADIEQEKDEGFLVMVYASQETFG 120
DB 61 SELVKIIRRLQLNPTQAFLLVNOHSMVSVTPDIADIEQEKDEGFLVMVYASQETFG 120

OY 121 F 121
DB 121 F 121

RESULT 3
US-08-805-117-3
Sequence 3, Application US/08805117
Patent No. 5955312

GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL MICROTUBULE-ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/805,117
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0211 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 142 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 455109
US-08-805-117-3

Query Match 83.4%; Score 521; DB 2; Length 142;
Best Local Similarity 82.5%; Pred. No. 1.7e-56;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

OY 1 MPSDRPFKORRSFADRCKEVOQIRDOHPSKIPVIERKGEKOLPVLDTKFLVDPDHVM 60
DB 1 MPSDRPFKORRSFADRCKEVOQIRDOHPSKIPVIERKGEKOLPVLDTKFLVDPDHVM 60
OY 61 SELVKIIRRLQLNPTQAFLLVNOHSMVSVTPDIADIEQEKDEGFLVMVYASQETFG 120
DB 61 SELVKIIRRLQLNPTQAFLLVNOHSMVSVTPDIADIEQEKDEGFLVMVYASQETFG 120

RESULT 4
US-09-199-838-3
Sequence 3, Application US/09199838
Patent No. 6280733
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL MICROTUBULE-ASSOCIATED PROTEIN

```

NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSER: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/199,838
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/805,117
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/PACKET NUMBER: PF-0211 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 142 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOP-LOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 455109
US-09-199-838-3

Query Match 83.4% Score 521; DB 3; Length 142;
Best Local Similarity 82.5%; Pred. NO. 1.7e-56;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 MPSDRPFKORSFADRCNEVOQIRDOHSKIPVITERTKGEQQLVLDKTKFLVPDHYNM 60
||||| ||||| |::| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MPSETFFKORSFEQRYVDVRLIREQHPKIPVITERTKGEQQLVLDKTKFLVPDHYNM 60

QY 61 SELKIIIRRLQNLNTOAFLLVNOHSMVSVSTPIADIJECKKDEGFGFYMYASQETFG 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 SELKIIIRRLQNLNNOAFLLVNOHSMVSVSTPISEVIESREDEGFLMYASQETFG 120

RESULT 5
US-09-708-725A-4
Sequence 4, Application US/09708725A
Patent No. 6489456
GENERAL INFORMATION:
APPLICANT: LADUNGA et al.
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
FILE REFERENCE: C1000842
CURRENT APPLICATION NUMBER: US/09/708,725A
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/243,428
PRIOR FILING DATE: 2000-10-27
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 4
LENGTH: 68
TYPE: PRT
ORGANISM: Human
US-09-708-725A-4

```

```

QY      42.7%   Score 697; DB 4; Length 68;
Query Match          42.7%; Score 697; DB 4; Length 68;
Best Local Similarity 85.2%; Pred. No. 1,1e-25;
Matches 52; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Db      119 F 119
       1
        61 F 61

RESULT 6
US-09-708-725A-2
Sequence 2, Application US/09708725A
Patent No. 6489456
GENERAL INFORMATION:
APPLICANT: LADUNGA et al.
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS.
TITLE OF INVENTION: NOCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS.
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: CLO00842
CURRENT APPLICATION NUMBER: US/09/7708, 725A
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/243,428
PRIOR FILING DATE: 2000-10-27
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 66
TYPE: PRT
ORGANISM: Human
US-09-708-725A-2

Query Match          42.1%; Score 263; DB 4; Length 66;
Best Local Similarity 85.2%; Pred. No. 3.4e-25;
Matches 52; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      60 MSEIIVKIRRLQLNPVDAFFILVNQHSMVSSTPIADIYEQEKDEMGFLMYVASQPEI 119
         |||||:|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1 MSSELIKIRRLQLNANQAFFILVNQHSMSVSSTPISSEYSEKDEMGFLMYVCASQPEI 60
                                         |||||  |||||  |||||  |||||  |||||

QY      120 G 120
         1
        61 G 61

RESULT 7
US-08-671-757A-9
Sequence 9, Application US/08671757A
Patent No. 6476213
GENERAL INFORMATION:
APPLICANT: Seierbaum, Sebastian
TITLE OF INVENTION: Cloning and Characterization Production
of Attagellate Strains
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flanagan, Henderson, Farlow, Garrett &
Danner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patentin Release #1.0, Version #1.30
APPLICATION NUMBER: US/08/671,757A

```

FILED DATE: 16-Aug-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 02366.0073-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 724 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-08-671-757A-9

Query Match
Best Local Similarity 12.4%; Score 77.5; DB 4; Length 724;
Matches 19; Conservative 25; Mismatches 42; Indels 1; Gaps 1;

DB 533 KEVQOIRDOHPSPKIVIIERYKGEKOLPVLDK-TKFLVDPDHVNMSELVKIIRRLQNP 76
QY 77 QAFLLVNOHMSVSVSTPIADYEOK 103
DB 593 SKSPMIEIHRASLARIAMITMYLDK 619

RESULT 8
US-08-369-822C-25
Sequence 25, Application US/08369822C
Patent No. 6015660
GENERAL INFORMATION:
APPLICANT: Lipkin, W. I.
APPLICANT: Biese, Thomas
APPLICANT: Kliche, Stefanie
APPLICANT: Schneider, Patrick A.
APPLICANT: Seltz, Lothar
APPLICANT: Schneemann, Anette
TITLE OF INVENTION: Borna Disease Viral Sequences,
TITLE OF INVENTION: Diagnostics and Therapeutics for Central Nervous
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESS: Fulbright & Jaworski, L.L.P.
STREET: 865 South Figueroa Street, 29th Floor
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90017-2571
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS NT-WORDEPERECT 8.0
SOFTWARE: ASCII (DOS) TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/369,822C
FILING DATE: 06-JAN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Churchill, Margaret A. (Ph.D.)
REGISTRATION NUMBER: 39,944
REFERENCE/DOCKET NUMBER: 1279-194XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213/680-4518
TELEFAX: 213/680-9200
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TOPOLOGY: linear

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-369-822C-25

Query Match
Best Local Similarity 11.2%; Score 70; DB 3; Length 527;
Matches 27; Conservative 19; Mismatches 46; Indels 36; Gaps 5;

DB 7 FKORSFADRCKEVOQIRD---QHPKIPVIER---YKGEKOLPVLDKTKFLVDPDHVN 59
DB 237 FESTALFGORCNEIFGKTFEFNMHP-----VLERCTIYGDPCPVADRMHROLQDHAD 291
QY 60 MSELVKIIRRLQNPQAFLLV-----NQHSVSVSTPIA 96
DB 292 SGIFLHNRGIE-GYCKLMTLISMALHIAVRGVRSAMVGDQDAIAVTSRPVVA 350
QY 97 DYEOKD 104
DB 351 QYKQKN 358

RESULT 9
US-08-582-776C-40
Sequence 40, Application US/08582776C
Patent No. 6077510
GENERAL INFORMATION:
APPLICANT: Lipkin, W. I.
APPLICANT: Biese, Thomas
APPLICANT: Kliche, Stefanie
APPLICANT: Schneider, Patrick A.
APPLICANT: Seltz, Lothar
APPLICANT: Schneemann, Anette
TITLE OF INVENTION: Borna Disease Viral Sequences,
TITLE OF INVENTION: Diagnostics and Therapeutics for Central Nervous
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESS: Fulbright & Jaworski, L.L.P.
STREET: 865 South Figueroa Street, 29th Floor
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90017-2576
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS NT
SOFTWARE: ASCII DOS TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/582,776C
FILING DATE: 04-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/369,822
FILING DATE: 06-JAN-1995
APPLICATION NUMBER: US 08/434,831
FILING DATE: 04-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Churchill, Margaret A.
REGISTRATION NUMBER: 39,944
REFERENCE/DOCKET NUMBER: 1279-194C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213/680-4518
TELEFAX: 213/680-9200
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: amino acid
TOPOLOGY: linear

Db 182 DTG 185

RESULT 12

US-09-398-395A-54
Sequence 54, Application US/09398395A
Patent No. 6468772
GENERAL INFORMATION:
APPLICANT: Chappell, Joseph
APPLICANT: No. 6468772, Joseph P.
APPLICANT: Starks, Courtney M.
APPLICANT: Manna, Kathleen R.
TITLE OF INVENTION: SYNTASES
FILE REFERENCE: 07678-025001
CURRENT APPLICATION NUMBER: US/09/398,395A
PRIOR FILING DATE: 1999-09-17
PRIOR APPLICATION NUMBER: 60/100,993
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/130,628
PRIOR FILING DATE: 1999-04-22
PRIOR APPLICATION NUMBER: 60/150,262
PRIOR FILING DATE: 1999-08-23
NUMBER OF SEQ ID NOS: 58
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 54
LENGTH: 590
TYPE: PR1
ORGANISM: Salvia officinalis
US-09-398-395A-54

Query Match

11.2%; Score 70; DB 4; Length 590;
Best Local Similarity 22.6%; Pred. No. 4;
Matches 28; Conservative 26; Mismatches 50; Indels 20; Gaps 4;

QY 6 PFQORSFADRCVEVOQIRQHSKIPVIERKGEKQLPVLDKTFVDPHNMSELYK 65
Db 73 PYKQORHFNQALIMQVRLKVKMEAI-----QQLLEIDDLQYLGLSYFFODEIKQ 125
QY 66 IIR-----RRLQNPQAFLLVNOHSMVSVTPADIYEOEKEDGFLYVYASQ 116
Db 126 ILSIHNEPRYFNNLDLYFALGFRLRQHG-FNSEVDYDCKRIKCSDF--FNANLQ 181
QY 117 ETFG 120
Db 182 DTG 185

RESULT 13

US-09-887-586A-54
Sequence 54, Application US/09887586A
Patent No. 6495354
GENERAL INFORMATION:
APPLICANT: Chappell, Joseph
APPLICANT: No. 6495354, Joseph P.
APPLICANT: Starks, Courtney M.
TITLE OF INVENTION: SYNTASES
FILE REFERENCE: 07678-025001
CURRENT APPLICATION NUMBER: US/09/887,586A
PRIOR FILING DATE: 2001-06-22
PRIOR APPLICATION NUMBER: 09/398,395
PRIOR FILING DATE: 1999-09-17
PRIOR APPLICATION NUMBER: 60/130,628
PRIOR FILING DATE: 1999-04-22
PRIOR APPLICATION NUMBER: 60/150,262
PRIOR FILING DATE: 1999-08-23
NUMBER OF SEQ ID NOS: 58
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 54
LENGTH: 590
TYPE: PR1
ORGANISM: Salvia officinalis

US-09-887-586A-54

Query Match

11.2%; Score 70; DB 4; Length 590;
Best Local Similarity 22.6%; Pred. No. 4;
Matches 28; Conservative 26; Mismatches 50; Indels 20; Gaps 4;

QY 6 PFQORSFADRCVEVOQIRQHSKIPVIERKGEKQLPVLDKTFVDPHNMSELYK 65
Db 73 PYKQORHFNQALIMQVRLKVKMEAI-----QQLLEIDDLQYLGLSYFFODEIKQ 125
QY 66 IIR-----RRLQNPQAFLLVNOHSMVSVTPADIYEOEKEDGFLYVYASQ 116
Db 126 ILSIHNEPRYFNNLDLYFALGFRLRQHG-FNSEVDYDCKRIKCSDF--FNANLQ 181
QY 117 ETFG 120
Db 182 DTG 185

RESULT 14

US-09-895-752-54
Sequence 54, Application US/09895752
Patent No. 6559297
GENERAL INFORMATION:
APPLICANT: Chappell, Joseph
APPLICANT: No. 6559297, Joseph P.
APPLICANT: Starks, Courtney M.
TITLE OF INVENTION: SYNTASES
FILE REFERENCE: 07678-025001
CURRENT APPLICATION NUMBER: US/09/895,752
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/398,395
PRIOR FILING DATE: 1999-09-17
PRIOR APPLICATION NUMBER: 60/100,993
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/130,628
PRIOR FILING DATE: 1999-04-22
PRIOR APPLICATION NUMBER: 60/150,262
PRIOR FILING DATE: 1999-08-23
NUMBER OF SEQ ID NOS: 58
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 54
LENGTH: 590
TYPE: PR1
ORGANISM: Salvia officinalis
US-09-895-752-54

Query Match

11.2%; Score 70; DB 4; Length 590;
Best Local Similarity 22.6%; Pred. No. 4;
Matches 28; Conservative 26; Mismatches 50; Indels 20; Gaps 4;

QY 6 PFQORSFADRCVEVOQIRQHSKIPVIERKGEKQLPVLDKTFVDPHNMSELYK 65
Db 73 PYKQORHFNQALIMQVRLKVKMEAI-----QQLLEIDDLQYLGLSYFFODEIKQ 125
QY 66 IIR-----RRLQNPQAFLLVNOHSMVSVTPADIYEOEKEDGFLYVYASQ 116
Db 126 ILSIHNEPRYFNNLDLYFALGFRLRQHG-FNSEVDYDCKRIKCSDF--FNANLQ 181
QY 117 ETFG 120
Db 182 DTG 185

RESULT 15

US-09-903-012B-54
Sequence 54, Application US/09903012B
Patent No. 6569656
GENERAL INFORMATION:
APPLICANT: Chappell, Joseph
APPLICANT: No. 6569656, Joseph P.
APPLICANT: Starks, Courtney M.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 10:15:34 : Search time 53 seconds

(Without alignments)
290.617 Million cell updates/sec

Title: US-09-904-603-1

Perfect score: 625
Sequence: 1 MSPDRPFKORRSPADRCREY.....EKDEDGLYVVASQETFCF 121

Scoring table: H10SUM62
Gapop 10.0 , Gapext 0.5

Searched: 479057 seqs, 127295195 residues

Total number of hits satisfying chosen parameters: 479057

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

```
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09C_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10C_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	625	100.0	121	11	US-09-904-603-1
2	521	83.4	142	11	US-09-904-603-3
3	521	83.4	142	12	US-10-205-194-71
4	521	83.4	171	9	US-09-925-301-1051
5	272	43.5	53	9	US-09-864-761-44587
6	267	42.7	68	15	US-10-277-156-4
7	263	42.1	66	15	US-10-277-156-2
8	248	39.7	117	15	US-10-205-342-23
9	71	11.4	501	9	US-09-815-242-5292
10	71	11.4	504	9	US-09-815-242-12503
11	70	11.2	590	9	US-09-887-585A-54
12	70	11.2	590	9	US-09-903-012-54
13	70	11.2	590	11	US-09-900-797-54
14	69.5	11.1	1233	10	US-09-738-626-4312
15	68.5	11.0	362	15	US-10-177-293-375

16	68.5	11.0	458	10	US-09-764-864-1164
17	68.5	11.0	578	15	US-10-137-416A-3
18	68.5	11.0	1360	15	US-10-137-416A-2
19	68	10.9	326	14	US-10-067-989-1
20	68	10.9	1616	11	US-09-820-843A-16
21	67	10.7	1236	11	US-10-128-714-3224
22	65	10.4	497	10	US-09-902-525-32
23	65	10.4	1814	9	US-09-920-552-103
24	65	10.4	3056	8	US-08-984-090-2
25	65	10.4	3056	11	US-09-825-476-3
26	65	10.4	3056	15	US-10-307-077-1
27	65	10.4	3056	15	US-10-175-225-2
28	65	10.4	3056	15	US-10-059-585-41
29	65	10.4	3057	15	US-10-175-225-3
30	64.5	10.3	324	10	US-05-981-876-181
31	64.5	10.3	324	11	US-09-148-545-181
32	64	10.2	317	15	US-10-023-282-504
33	64	10.2	1008	10	US-09-866-582-39
34	63.5	10.2	83	9	US-09-867-550-498
35	63	10.1	220	9	US-09-815-242-10585
36	63	10.1	1718	15	US-10-223-070-21
37	62.5	10.0	190	14	US-10-108-605-167
38	62.5	10.0	267	15	US-10-189-346-89
39	62.5	10.0	283	14	US-10-042-417-50
40	62.5	10.0	586	11	US-09-878-131-2
41	62.5	10.0	1257	15	US-10-128-714-8224
42	62.5	10.0	26926	10	US-09-759-508B-2
43	62	9.9	220	9	US-09-938-803-4
44	62	9.9	335	15	US-10-156-761-7713
45	62	9.9	384	10	US-09-801-876B-6

ALIGNMENTS

```
RESULT 1
US-09-904-603-1
: Sequence 1, Application US/09904603
: Publication NO. US20030095612A1
:
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: Goll, Surya K.
: TITLE OF INVENTION: NOVEL MICROBUTYLE ASSOCIATED PROTEIN
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSER: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/904,603
: FILING DATE: 12-Jul-2001
: CLASSIFICATION: <unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/805,117
: FILING DATE: <unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0211 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4196
:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
```

```

LENGTH: 121 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: THYRNOT03
CLONE: 144378
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-904-603-1

Query Match      100.0%; Score 625; DB 11; Length 121;
Best Local Similarity 100.0%; Prid. No. 1,2e-64;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MPSDRPFKQRSFADRCKEVQGIIDQHPKIPVITIERYKGEKOLPVLDTKTFPLPDHYNM 60
        1 MPSDRPFKQRSFADRCKEVQGIIDQHPKIPVITIERKGEKOLPVLDTKTFPLPDHYNM 60
        1 MPSDRPFKQRSFADRCKEVQGIIDQHPKIPVITIERKGEKOLPVLDTKTFPLPDHYNM 60

QY      61 SELVKIIRRLQLNPTQAFLLVNMHSMVSVSTPIADIYEOEKDEGFLYMYVASQETFG 120
        61 SELVKIIRRLQLNPTQAFLLVNMHSMVSVSTPIADIYEOEKDEGFLYMYVASQETFG 120
        61 SELVKIIRRLQLNPTQAFLLVNMHSMVSVSTPIADIYEOEKDEGFLYMYVASQETFG 120

QY      121 F 121
        121 F 121
        121 F 121

Db      121 F 121

RESULT 2
US-09-904-603-3
Sequence 3, Application US/09904603
Publication No. US20030099612A1
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
Goll, Surya K.
TITLE OF INVENTION: NOVEL MICROFIBULE-ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/904,603
FILING DATE: 12-Jul-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/805,117
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0211_US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <unknown>
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 142 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 455109
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

```

US-09-904-603-3
Query Match      83.4% Score 521; DB 11; Length 142:
Best Local Similarity 82.5%; Pred. No. 1,5e-52;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY      1 MPSDRPFKORRSFADRCKEYVQOIRDOHPKIPVIIEERYKSGEKQLPVLDKTKFLVPDHHVM 60
        |||:||||||| | :||: ||:||||||| |||:||||||| |||:||||||| |||:|||||||
Db       1 MPSKTEFKORRSFQQRVEDVRLIREQHPTKIPVIIEERYKSGEKQLPVLDKTKFLVPDHHVM 60
        |||:||||||| | :||: ||:||||||| |||:||||||| |||:||||||| |||:|||||||

QY      61 SELVKIIRRRLOLNPOTAFLLVNQSHMSVSPTPIADIYEDEKEDGFLVMVVASOETFG 120
        |||:||||||| | :||: ||:||||||| |||:||||||| |||:||||||| |||:|||||||
Db       61 SELKIIRRRLOLANAQAFLLVNGHSMVSPTPISEYSESRDEDGFLVMVVASOETFG 120
        |||:||||||| | :||: ||:||||||| |||:||||||| |||:||||||| |||:|||||||

RESULT 3
US-10-205-194-71
Sequence 71, Application US/10205194
Publication No. US20030134301A1
GENERAL INFORMATION:
APPLICANT: Warner-Lambert Company
APPLICANT: Lee, Kevin
APPLICANT: Dixon, Alister
APPLICANT: Brooksbank, Robert
APPLICANT: Plimock, Robert
TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
FILE REFERENCE: WL-A-018201
CURRENT APPLICATION NUMBER: US/10/205,194
CURRENT FILING DATE: 5200-07-24
PRIOR APPLICATION NUMBER: GB 0118354.0
PRIOR FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 177
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 71
LENGTH: 142
TYPE: PRT
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Light chain 3 subunit of microtubule-associated proteins 1A and
US-10-205-194-71

Query Match      83.4% Score 521; DB 12; Length 142:
Best Local Similarity 82.5%; Pred. No. 1,5e-52;
Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY      1 MPSDRPFKORRSFADRCKEYVQOIRDOHPKIPVIIEERYKSGEKQLPVLDKTKFLVPDHHVM 60
        |||:||||||| | :||: ||:||||||| |||:||||||| |||:||||||| |||:|||||||
Db       1 MPSKTEFKORRSFQQRVEDVRLIREQHPTKIPVIIEERYKSGEKQLPVLDKTKFLVPDHHVM 60
        |||:||||||| | :||: ||:||||||| |||:||||||| |||:||||||| |||:|||||||

QY      61 SELVKIIRRRLOLNPOTAFLLVNQSHMSVSPTPIADIYEDEKEDGFLVMVVASOETFG 120
        |||:||||||| | :||: ||:||||||| |||:||||||| |||:||||||| |||:|||||||
Db       61 SELKIIRRRLOLANAQAFLLVNGHSMVSPTPISEYSESRDEDGFLVMVVASOETFG 120
        |||:||||||| | :||: ||:||||||| |||:||||||| |||:||||||| |||:|||||||

RESULT 4
US-09-925-301-1051
Sequence 1051, Application US/09925301
Patent No. US2002052308A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1051
LENGTH: 171
TYPE: PRT

```

```

1 ORGANISM: Homo sapiens
2 US-09-925-301-1051
3
4 Query Match      83.4%; Score 521; DB 9; Length 171;
5 Best local similarity 82.5%; Pred. No. 2e-52;
6 Matches 99; Conservative 12; Mismatches 9; Indels 0; Gaps 0;
7
8 QY          1 MSPDPPFRRKFEVADRCQVOOIRDOHPKIPYIERKYKEKKOLPYLDTKFLPDPHVM 60
9             |||::|||::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
10 Db          47 MPSEFTFRORIRFEQVRVEDVRLIREOHPIKIPYIERKYKEKKOLPYLDTKFLPDPHVM 106
11             |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
12 QY          61 SELVKIIRRIQLNPTQAEFLLVNCHSMVSVPSTADIYEQEKDEDDGFLYMVASQETFG 120
13             |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
14 Db          107 SELIKIIRRIQLNANQAFFLLVNGHSMSVSPSTIEVESEKDEDGFLYMVASQETFG 166
15             |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
16
17 RESULT 5
18 US-09-864-761-44587
19 Sequence 44587, Application US/09864761
20 Patent No. US20020048763A1
21 GENERAL INFORMATION:
22 APPLICANT: Peno, Sharon G.
23 APPLICANT: Kank, David R.
24 APPLICANT: Hanzel, David K.
25 APPLICANT: Chen, Wensheng
26 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
27 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
28 FILE REFERENCE: Aeomica-X-1
29 CURRENT APPLICATION NUMBER: US/09/864,761
30 CURRENT FILING DATE: 2001-05-23
31 PRIOR APPLICATION NUMBER: US 60/180,312
32 PRIOR FILING DATE: 2000-02-04
33 PRIOR APPLICATION NUMBER: US 60/207,456
34 PRIOR FILING DATE: 2000-05-26
35 PRIOR APPLICATION NUMBER: US 09/632,366
36 PRIOR FILING DATE: 2000-08-03
37 PRIOR APPLICATION NUMBER: GB 24263.6
38 PRIOR FILING DATE: 2000-10-04
39 PRIOR APPLICATION NUMBER: US 60/236,359
40 PRIOR FILING DATE: 2000-09-27
41 PRIOR APPLICATION NUMBER: PCT/US01/00666
42 PRIOR FILING DATE: 2001-01-30
43 PRIOR APPLICATION NUMBER: PCT/US01/00667
44 PRIOR FILING DATE: 2001-01-30
45 PRIOR APPLICATION NUMBER: PCT/US01/00664
46 PRIOR FILING DATE: 2001-01-30
47 PRIOR APPLICATION NUMBER: PCT/US01/00669
48 PRIOR FILING DATE: 2001-01-30
49 PRIOR APPLICATION NUMBER: PCT/US01/00665
50 PRIOR FILING DATE: 2001-01-30
51 PRIOR APPLICATION NUMBER: PCT/US01/00668
52 PRIOR FILING DATE: 2001-01-30
53 PRIOR APPLICATION NUMBER: PCT/US01/00663
54 PRIOR FILING DATE: 2001-01-30
55 PRIOR APPLICATION NUMBER: PCT/US01/00662
56 PRIOR FILING DATE: 2001-01-30
57 PRIOR APPLICATION NUMBER: PCT/US01/00661
58 PRIOR FILING DATE: 2001-01-30
59 PRIOR APPLICATION NUMBER: PCT/US01/00670
60 PRIOR FILING DATE: 2001-01-30
61 PRIOR APPLICATION NUMBER: US 60/234,687
62 PRIOR FILING DATE: 2000-09-21
63 PRIOR APPLICATION NUMBER: US 09/608,408
64 PRIOR FILING DATE: 2000-06-30
65 PRIOR APPLICATION NUMBER: US 09/774,203
66 PRIOR FILING DATE: 2001-01-29
67 NUMBER OF SEQ ID NOS: 49117
68 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
69 SEQ ID NO 44587
70 LENGTH: 53
71 TYPE: PRT
72 ORGANISM: Homo sapiens
73 FEATURE:

```

```

1  OTHER INFORMATION: MAP TO AL116520.12
2  OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.67
3  OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
4  OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.66
5  OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.66
6  OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.65
7  OTHER INFORMATION: EST_HUMAN HIT: HE712850.1, EVALU6 5.00e-24
8  OTHER INFORMATION: SWISSPROT HIT: Q62625, EVALU6 3.00e-20
9  US-09-864-761-44587
10
11 Query Match          43.5%; Score 272; DB 9; Length 54;
12 Best Local Similarity 100.0%; Pred. No. 2,6e-24;
13 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
14
15 QY      69  RLQIINTQAFFLLVNHSHMSVSTPIADIYQEKDEKDEGFLYMYVASQETFGF 121
16      |||
17      1  RLQIINTQAFFLLVNHSHMSVSTPIADIYQEKDEKDEGFLYMYVASQETFGF 54
18
19 RESULT 6
20 US-10-277-156-4
21 Sequence 4, Application US/10277156
22 Publication No. US20030054399A1
23 GENERAL INFORMATION:
24 APPLICANT: LADUNGA et al.
25 TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS.
26 TITLE OF INVENTION: NUCLEIC ACID MULTIPLES ENCODING HUMAN SECRETED PROTEINS.
27 TITLE OF INVENTION: USES THEREOF
28 FILE REFERENCE: CLO00842 DIV
29 CURRENT APPLICATION NUMBER: US/10/277,156
30 CURRENT FILING DATE: 2002-10-22
31 PRIOR APPLICATION NUMBER: 09/7708,725
32 PRIOR FILING DATE: 2000-11-09
33 PRIOR APPLICATION NUMBER: 60/243,428
34 PRIOR FILING DATE: 2000-10-27
35 NUMBER OF SEQ ID NOS: 4
36 SOFTWARE: FastSeq for Windows Version 4.0
37 SEQ ID NO 4
38 LENGTH: 68
39 TYPE: PRT
40 ORGANISM: Human
41 US-10-277-156-4
42
43 Query Match          42.7%; Score 267; DB 15; Length 68;
44 Best Local Similarity 85.2%; Pred. No. 1,4e-23;
45 Matches 52; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
46
47 QY      59  NMSLVKIKRKRLQIINTQAFFLLVNHSHMSVSTPIADIYQEKDEKDEGFLYMYVASQET 118
48      |||
49      1  NMSLVKIKRKRLQIINTQAFFLLVNHSHMSVSTPIADIYQEKDEKDEGFLYMYVASQET 60
50
51 QY      119  F 119
52      |
53      61  F 61
54
55 RESULT 7
56 US-10-277-156-2
57 Sequence 2, Application US/10277156
58 Publication No. US20030054399A1
59 GENERAL INFORMATION:
60 APPLICANT: LADUNGA et al.
61 TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS.
62 TITLE OF INVENTION: NUCLEIC ACID MULTIPLES ENCODING HUMAN SECRETED PROTEINS.
63 TITLE OF INVENTION: USES THEREOF
64 FILE REFERENCE: CLO00842 DIV
65 CURRENT APPLICATION NUMBER: US/10/277,156
66 CURRENT FILING DATE: 2002-10-22
67 PRIOR APPLICATION NUMBER: 09/7708,725
68 PRIOR FILING DATE: 2000-11-09
69 PRIOR APPLICATION NUMBER: 60/243,428
70 PRIOR FILING DATE: 2000-10-27
71 NUMBER OF SEQ ID NOS: 4
72

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Human
US-10-277-156-2

Query Match
Best Local Similarity 42.1%; Score 263; DB 15; Length 66;
Matches 52; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 60 MSELVKIRRLQINPQAFLLVNOHSMVSVSTPIADIYEQEKDEGFLYVYASQETP 119
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MSELIKIRRLQINANQAFLLVNGHSMVSVSTPISEYSEKDEGFLYVYASQETP 60

OY 120 G 120
Db 61 G 61

RESULT 8
US-10-205-342-23
; Sequence 23, Application US/10205342
; Publication No. US20030108906A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alstair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Plincock, Robert
; TITLE OF INVENTION: Identification and use of molecules implicated in pain
; FILE REFERENCE: WU-A-018198
; CURRENT APPLICATION NUMBER: US/10/205,342
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Protein: Ganglioside expression factor 2
US-10-205-342-23

Query Match
Best Local Similarity 39.7%; Score 248; DB 15; Length 117;
Matches 48; Conservative 29; Mismatches 36; Indels 2; Gaps 2;

OY 7 FKQRRSFADRCKEVOQIRDOHPKPIYIERKGEKQLPYLDKTKFLVDPHVNMSLVKI 66
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 5 FKEDHSLEHRCVESAKIRAKYPRDPIYIEKVSQ-SQIVIDKRRKLVPSDITVAQFMWI 63

OY 67 IRRRLQINPQAFLLVNOHSMVSVSTPIADIYEQEKDEGFLYVYASQETP 121
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 64 IRRRIQPSKRAIFLFVDK-TVPOSSLTQMQLYEKEDGFLYVYASGNTGTF 117

RESULT 9
US-09-815-242-5292
; Sequence 5292, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlssen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
```

```

; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5292
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(501)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-5292

Query Match
Best Local Similarity 11.4%; Score 71; DB 9; Length 501;
Matches 26; Conservative 32; Mismatches 47; Indels 26; Gaps 6;

OY 8 KORRSFADRCKEVOQIRD-----OHPSKIPYIERKGEKQL--PYLDKTKFLVDPH 57
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 231 KNRMEETVYKDLTTIDEEFTDYGKORPEILPTDNNENKALAKSLDLNITSNS 290

OY 58 VNMSLVKIIIR-RLQINPQAFLLVNOHSMVSVST-----PIADIYEQEKDEGFL 108
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 291 SQKKTVDKVVKKPKDQSTGKKEF--VIKEPTKPSIPTYTEIRKKPAPSIKQELKEK--- 345

OY 109 LYNVYASQETP 119
Db 346 --IIVSSPDRF 354

RESULT 10
US-09-815-242-12503
; Sequence 12503, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlssen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
```

```

: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO: 12503
: LENGTH: 504
: TYPE: PRI
: ORGANISM: Staphylococcus aureus
US-09-815-242-12503

Query Match          11.4%  Score 71;  DB 9;  Length 504;
Best Local Similarity 19.8%;  Pred. No. 9.6;
Matches 26;  Conservative 32;  Mismatches 47;  Indels 26;  Gaps 6;

UY      8  KORSEFADCKEVOQIRD-----OHPSKIPVILIERYKGEKOL--PVLDTKFLVPDH 57
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      249  KNRMEFEITYKDELTINDEFDTDGKORPSEIPTLTNNENKALASRLQITSTNSNS 308
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
UY      58  VNNSVLKRIIR-RLQINPTQAFLLVNOHSMVSVT-----PIADIYEDKEDGDF 108
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      109  SOKKTVDKVKKPKDQSTGNKEF--VIKEPTKPSIPTYEITKPKMPSIKQETIKER--- 363
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
UY      109  LYWYASQETP 119
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      364  --IIYSSPQRF 372
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 11
US-09-887-586A-54
: Sequence 54, Application US/09887586A
: Patent No. US20020094556A1
: GENERAL INFORMATION:
: APPLICANT: Chappell, Joseph
: APPLICANT: No. US20020094556A11, Joseph P.
: APPLICANT: Starks, Courtney M.
: APPLICANT: Manna, Kathleen R.
: TITLE OF INVENTION: SYNTHASES
: FILE REFERENCE: 07678-025001
: CURRENT APPLICATION NUMBER: US/09/887,586A
: PRIOR FILING DATE: 2001-06-22
: PRIOR APPLICATION NUMBER: 09/398,395
: PRIOR FILING DATE: 1999-09-17
: PRIOR APPLICATION NUMBER: 60/130,628
: PRIOR FILING DATE: 1999-04-22
: PRIOR APPLICATION NUMBER: 60/150,262
: PRIOR FILING DATE: 1999-08-23
: NUMBER OF SEQ ID NOS: 58
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 54
: LENGTH: 590
: TYPE: PRI
: ORGANISM: Salvia officinalis
US-09-887-586A-54

Query Match          11.2%  Score 70;  DB 9;  Length 590;
Best Local Similarity 22.6%;  Pred. No. 16;
Matches 28;  Conservative 26;  Mismatches 50;  Indels 20;  Gaps 4;
```

```

RESULT 12
US-09-903-012-54
: Sequence 54, Application US/09903012
: Patent No. US20020094557A1
: GENERAL INFORMATION:
: APPLICANT: Chappell, Joseph
: APPLICANT: No. US20020094557A11, Joseph P.
: APPLICANT: Starks, Courtney M.
: APPLICANT: Manna, Kathleen R.
: TITLE OF INVENTION: SYNTHASES
: FILE REFERENCE: 07678-025001
: CURRENT APPLICATION NUMBER: US/09/903,012
: PRIOR FILING DATE: 2001-07-11
: PRIOR APPLICATION NUMBER: 09/498,495
: PRIOR FILING DATE: 1999-09-17
: PRIOR APPLICATION NUMBER: 60/100,993
: PRIOR FILING DATE: 1998-09-18
: PRIOR APPLICATION NUMBER: 60/140,628
: PRIOR FILING DATE: 1999-04-22
: PRIOR APPLICATION NUMBER: 60/150,262
: PRIOR FILING DATE: 1999-08-23
: NUMBER OF SEQ ID NOS: 58
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 54
: LENGTH: 590
: TYPE: PRI
: ORGANISM: Salvia officinalis
US-09-903-012-54

Query Match          11.2%  Score 70;  DB 9;  Length 590;
Best Local Similarity 22.6%;  Pred. No. 16;
Matches 28;  Conservative 26;  Mismatches 50;  Indels 20;  Gaps 4;
```

```

UY      6  PFKORSEFADCKEVOQIRDQHSKIPVILIERYKGEKOLPVLDTKFLVPDHVNMSELYK 55
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      73  PYKQHRFNQAEILMQVRLMLKVKMEAI-----QOLELIDLOYLGLSYFFQDEIKQ 125
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
UY      66  IIR-----RRLQINPTQAFLLVNOHSMVSVTPIADIYEDKEDGDFLYWYASQ 116
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      126  ILSSTHNEPRYFNHNDLYFTALGFRILRQHG- NVSDVDFGCFIEKCSD---FNANLAO 181
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
UY      117  EFTG 120
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      182  DTG 185
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 13
US-09-900-797-54
: Sequence 54, Application US/09900797
: Publication No. US20030087406A1
: GENERAL INFORMATION:
: APPLICANT: Chappell, Joseph
: APPLICANT: No. US20030087406A11, Joseph P.
: APPLICANT: Starks, Courtney M.
: APPLICANT: Manna, Kathleen R.
: TITLE OF INVENTION: SYNTHASES
: FILE REFERENCE: 07678-025001
: CURRENT APPLICATION NUMBER: US/09/900,797
: PRIOR FILING DATE: 2001-07-06
: PRIOR APPLICATION NUMBER: US/09/398,395
: PRIOR FILING DATE: 1999-09-17
: PRIOR APPLICATION NUMBER: 60/130,628
: PRIOR FILING DATE: 1999-04-22
: PRIOR APPLICATION NUMBER: 60/150,262
: PRIOR FILING DATE: 1999-08-23
: NUMBER OF SEQ ID NOS: 58
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 54
: LENGTH: 590
: TYPE: PRI
: ORGANISM: Salvia officinalis
```

US-09-900-797-54

Query Match 11.2%; Score 70; DB 11; Length 590;
Best Local Similarity 22.6%; Pred. No. 16;
Matches 28; Conservative 26; Mismatches 50; Indels 20; Gaps 4;

OY 6 PFKRRSFADRCKEVQIRDOHPSPKIPIIERYGKQOLPYLDKTRFLVDPHVNSELVK 65
DB 73 PYKEQRHFRHROAEILMGOVMRLKVKMEAI-----QOLELIDDLQYLGISYFODEIKQ 125
OY 66 IIR-----RRLQNPQAFLLVNHQSMVSTPIADIYEDKEKDEGFLVNVYSQ 116
DB 126 ILSSTHNEPRYFHNNDLFTALGFRLKQHG-FNSEVDFCCFKIEKCD--FNANLAQ 181
OY 117 ETPFG 120
DB 182 DTGK 185

RESULT 14

US-09-738-626-4312
Sequence 4312, Application US/09738626
Publication No. US20020197605A1

GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 4312
LENGTH: 1233
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-4312

Query Match 11.1%; Score 69.5; DB 10; Length 1233;
Best Local Similarity 18.8%; Pred. No. 48;
Matches 34; Conservative 19; Mismatches 39; Indels 89; Gaps 5;

OY 1 MPSDRPFQRRSFADRCKEVQOIRDO----- 26
DB 1 MTSRRLKQERSFADDDLODKLTNDQLFTNAKLQARISGNDGKKTTRPTPLALDQ 60
OY 27 -----HPSKTVITIERKGEKQPLVLDKTKFL----- 53
DB 61 LTVBEYETITAILVEAVGNGSKPAILKDLFEYPLVFLALSGTAMLDAQEGFWPAFWK 120
OY 54 -----VDPHVN-----MSELVKIIRRRLONPQAFLL-----LVNQHSMVSVSTPI 95
DB 121 RTQVSVEPEHYDAIRKELVNSIRK-----NGLETSLADLNRRREVGLQHSGLSADML 176
OY 96 A 96
DB 177 A 177

RESULT 15

US-10-177-293-375

Sequence 375, Application US/10177293
Publication No. US20030124128A1
GENERAL INFORMATION:

APPLICANT: Lillie, James
APPLICANT: Glatt, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Gannavarpu, Manjula
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Mertens, Maureen
APPLICANT: Myer, Vic
APPLICANT: Wang, Youzhen
APPLICANT: Xu, Yongyao
APPLICANT: Hoersch, Sebastian
APPLICANT: Monahan, John
APPLICANT: Meyers, Rachel E.
APPLICANT: Bast Jr., Robert C.
APPLICANT: Hortobagyi, Gabriel N.
APPLICANT: Puzstai, Lajos
APPLICANT: Meric, Funda
APPLICANT: Sahlin, Aysegül
APPLICANT: Mills, Gordon B.
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT
FILE REFERENCE: MRI-038
CURRENT APPLICATION NUMBER: US/10/177,293
PRIOR FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: US 60/299,887
PRIOR FILING DATE: 2001-06-21
PRIOR APPLICATION NUMBER: US 60/301,572
PRIOR FILING DATE: 2001-06-27
PRIOR APPLICATION NUMBER: US 60/306,501
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/325,002
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/362,585
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 60/xxx,xxx
PRIOR FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 506
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 375
LENGTH: 362
TYPE: PRT
ORGANISM: Homo sapiens
US-10-177-293-375

Query Match 11.0%; Score 68.5; DB 15; Length 362;
Best Local Similarity 29.3%; Pred. No. 12;
Matches 24; Conservative 12; Mismatches 37; Indels 9; Gaps 3;

OY 41 EKOLPYLDK-TKFLVDPHVNSELVKIIRRRLONPQAFLLVNHQSMVSVSTPIADIY 99
DB 285 EKALAILSQPTPSLVVHERLKNLTKTVKKSQ---NWNIFOLENLVAIVISQC-----IY 336
OY 100 EOKEKDEGFLVNVASQETGCF 121
DB 337 RHRKDHTKSLQKMEQEVENF 358

Search completed: August 11, 2003, 10:24:37
Job time : 53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: August 11, 2003, 10:09:23 ; Search time 38 Seconds

(without alignments)
306.221 Million cell updates/sec

Title: US-09-904-603-1

Perfect score: 625

Sequence: 1 MPSDRPKQRKRSFADCKEY.....EKDEGFLYMYVASQETFCF 121

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

SUMMARIES

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	521	83.4	142	2	A53624
2	381.5	61.0	130	2	T27920
3	227	36.3	120	2	H84470
4	225	35.0	122	2	C84887
5	224	35.8	121	2	T40818
6	219	35.0	107	2	T49105
7	219	35.0	122	2	C85058
8	217	34.7	121	2	T47875
9	216	34.6	117	2	S45432
10	203	32.5	117	2	JC7698
11	201.5	32.2	149	2	T02148
12	194	31.0	123	2	T15740
13	140	22.4	165	2	B71432
14	85	13.6	841	2	B71212
15	85	13.6	1210	2	H88451
16	79.5	12.7	282	2	T27449
17	77.5	12.4	500	2	E96735
18	77.5	12.4	724	2	A49217
19	77.5	12.4	724	2	C81361
20	75	12.3	627	2	T49837
21	75	12.0	592	2	T50154
22	74.5	11.9	660	2	A81704
23	74	11.8	790	2	T19683
24	73	11.7	260	2	S60480
25	73	11.7	746	2	T43648
26	72.5	11.6	224	2	F69937
27	72.5	11.6	682	2	T23813
28	72	11.5	483	2	F64204
29	72	11.5	1046	1	GNMVC

30 72 11.5 1761 2 114289 DNA (cytosine-rich)
31 71 11.4 392 2 659647 glycine-rich repeat
32 71 11.4 558 2 164509 hypothetical protein
33 71 11.4 558 2 164509 hypothetical protein
34 71 11.4 2048 1 ZLN2SE genome-folyprotein
35 70.5 11.3 278 2 C64393 hypothetical protein
36 70 11.2 132 2 T50108 yeast Acp12 protein
37 70 11.2 192 2 T12775 hypothetical protein
38 70 11.2 1094 2 S49418 protein kinase
39 70 11.2 1272 2 T130248 fragile X mental
40 70 11.2 2226 1 ZLN2SV genome-folyprotein
41 69.5 11.1 262 2 140221 divin protein
42 69 11.0 416 2 116205 hypothetical protein
43 69 11.0 704 2 S13263 transfer RNA
44 68.5 11.0 453 2 B69504 seryl-transfer
45 68 10.9 242 2 C95219 conserved hypothet

ALIGNMENTS

RESULT 1

A53624 microtubule-associated protein 1 light chain 3 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 24 Sep-1999

C:Accession: A53624; S45180

R:Man, S.S.; Hamada, J.A.

J. Biol. Chem. 269, 11492-11497, 1994

A:Title: Molecular characterization of light chain 3. A microtubule-binding subunit of

A:Reference number: A53624; MIM:420931; PMID:7908909

A:Accession: A53624

A:Molecule type: mRNA

A:Residues: 1-142 <MAN>

A:Cross-references: EMBL:005784; NID:0455108; P1DN:AAA20645.1; P11:0455109

C:Superfamily: hypothetical protein YH078c

Query Match 83.4% Score 521 DB 2 Length 142

Best Local Similarity 82.5% Pred. No. 6,7e+44

Matches 99; Conservative 12; Mismatches 9; Indels 0 Gaps 0

OY 1 MPSDRPKQRKRSFADCKEYVGLRDNRSKIVLPRKQKQYRVDKTKLTVRNM 50

Db 1 MSEKTFKQVRSEGVETLRLRDNRIKTVLTERKQKQYRVDKTKLTVRNM 60

OY 61 SEVKTIRKRLQINPTAFELLVNDSMYSSTRADYVDEKDEYEFYMYVASQETFG 120

Db 61 SEVKTIRKRLQINPTAFELLVNDSMYSSTRADYVDEKDEYEFYMYVASQETFG 120

RESULT 2

T27920 hypothetical protein ZK593.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000

C:Accession: T27920

R:McMurray, A.

submitted to the EMBL Data Library, February 1996

A:Reference number: Z20440

A:Accession: T27920

A>Status: preliminary; translated from GH/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-130 <WIL>

A:Cross-references: EMBL:Z69485; P1DN:AA34421.1; USDB:GN00022; CESP:ZK593.6

A:Experimental source: clone ZK934

C:Genetics:

A:Gene: CESP:ZK593.6

A:Map position: 4

A:Introns: 24/1: 86/2

C:Superfamily: hypothetical protein YH078c

Query Match 61.0% Score 381.5 DB 2 Length 142

Best Local Similarity 58.3% Pred. No. 2.8e+30

Matches 70: Conservative 27; Mismatches 20; Indels 3; Gaps 1;

QY 1 MPSDRPFKORSPADCKEVOQIRDOHPSKIPVIIRYKGEKQLPVLDTKFLVDPHVM 60
 14 VPS---FKERPRPHERKQVDEIRSOQPNKVPVILRPFGEKSLPLMDCKFLVPEHIV 70

Db 61 SELVITIRRLQNPQAFLVNOHSMVSTPIADIYEOEKDEGFLYVYASQETFG 120
 71 AELMSIVRRRLQLHPQDAFLVNERSMVSNMSMNLVYQERDPGFFVYVMTSQPAFG 130

RESULT 3
 H84470
 Probable microtubule-associated protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: H84470
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: H84470
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-120 <STO>
 A:Cross-references: GB:AE002093; NID:94581162; PIDN:AMD24645.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g05630
 A:Map position: 2
 C:Superfamily: hypothetical protein YBL078c

Query Match 36.3%; Score 227; DB 2; Length 120;
 Best Local Similarity 37.7%; Pred. No. 3.5e-15;
 Matches 43; Conservative 32; Mismatches 37; Indels 2; Gaps 2;

QY 7 FKORSPADCKEVOQIRDOHPSKIPVIIRYKGEKQLPVLDTKFLVDPHVMSELVKI 66
 6 FKHEPFLKROAEARIRKIPDRIPVIEPAE-KSDVPDIDRRKILVPAALTYGQFPYV 64

Db 67 IRRRLQNPQAFLVNOHSMVSTPIADIYEOEKDEGFLYVYASQETFG 120
 65 VKRIKLSPEKAIFIFV-KNILPPTAALMSAIEEHKDEGFLYVYSGENTFG 117

RESULT 4
 C84887
 Probable microtubule-associated protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: C84887
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: C84887
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-122 <STO>
 A:Cross-references: GB:AE002093; NID:96598366; PIDN:AAFI8602.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g45170
 A:Map position: 2
 C:Superfamily: hypothetical protein YBL078c

Query Match 36.0%; Score 225; DB 2; Length 122;
 Best Local Similarity 39.5%; Pred. No. 5.5e-15;
 Matches 47; Conservative 29; Mismatches 31; Indels 12; Gaps 4;

QY 7 FKORSPADCKEVOQIRDOHPSKIPVIIRYKGEKQLPVLDTKFLVDPHVMSELVK 65

Db 7 FKMDDFEKKRAEAGRIKIPDRIPVIE---KAEKSVPIIDKKKYLVPDLSLVGQFPY 64

QY 66 IRRRLQNPQAFLVNOHSMVSTPIADIYEOEKDEGFLYVYASQETFG 122
 65 VKRIKLSPEKAIFIFVD-----NVLPTGELMSVYEDKKDEGFLYVYSGENTFG 116

RESULT 5
 T40818
 Probable autophagy protein - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T40818
 R:Beck, A.; Reinhardt, R.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
 Submitted to the EMBL Data Library, October 1998
 A:Reference number: 221949
 A:Accession: T40818
 A:Status: preliminary
 A:Molecule type: translated from GB/EMBL/DBJ
 A:Residues: 1-121 <BEC>
 A:Cross-references: EMBL:AL032684; PIDN:CAN21809.1; GSPDB:GN00067; SPDB:SPBP887.24
 A:Experimental source: Strain 972h-, clone p1 p887
 C:Genetics:
 A:Gene: SPDB:SPBP887.24c
 A:Map position: 2
 A:Introns: 12/1; 30/3
 C:Superfamily: hypothetical protein YBL078c

Query Match 35.8%; Score 224; DB 2; Length 121;
 Best Local Similarity 37.7%; Pred. No. 6.9e-15;
 Matches 43; Conservative 31; Mismatches 38; Indels 2; Gaps 2;

QY 7 FKORSPADCKEVOQIRDOHPSKIPVIIRYKGEKQLPVLDTKFLVDPHVMSELVKI 66
 5 FKDFSEKKTTSQRIKIPDRIPVIECKV-DKSDIAIDKKKYLVPDLSLVGQFPYV 63

Db 67 IRRRLQNPQAFLVNOHSMVSTPIADIYEOEKDEGFLYVYASQETFG 120
 64 IRRKIKLSPEKAIFIFIDE-ILPPTAALMSVIEEHKSEDEGFLYVYSGENTFG 116

RESULT 6
 T49105
 Symbols-related like protein - Arabidopsis thaliana
 N:Alternate names: protein AT4g21980
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Sep-2000
 R:Bevan, M.; Wedler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.
 submitted to the Protein Sequence Database, May 2000
 A:Reference number: 225016
 A:Accession: T49105
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-107 <BEV>
 A:Cross-references: EMBL:AL022140; GSPDB:GN00062; ATSP:AT4g21980
 A:Experimental source: cultivar Columbia; BAC clone PIN20
 C:Genetics:
 A:Gene: ATSP:AT4g21980
 A:Map position: 4
 A:Introns: 16/3; 34/2; 74/1
 C:Superfamily: hypothetical protein YBL078c

Query Match 35.0%; Score 219; DB 2; Length 107;
 Best Local Similarity 39.2%; Pred. No. 1.8e-14;
 Matches 40; Conservative 31; Mismatches 29; Indels 2; Gaps 2;

QY 19 EVQOIRDOHPSKIPVIIRYKGEKQLPVLDTKFLVDPHVMSELVKIIRRLQNPQA 78
 3 ESSRIREKYPRIPIYIEK-AGQSDVDPIDKKKYLVPADLVGQFPYVVKRIKLSA 61

Db 79 FFLVNOHSMVSTPIADIYEOEKDEGFLYVYASQETFG 120

Db 62 IFVFEV-KNLTPTALMSATIEFHKDEDEGLYMTYSGENTFG 102

RESULT 7

C85058

Probable symbiosis-related protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001

C:Accession: C85058

R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring

Nature 402: 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: A85001; MUID:20083488; PMID:10617158

A:Accession: C85058

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-122 <STD>

A:Cross-references: GB:NC_001268; NID:g7267220; PIDN:CA80827.1; GSPDB:GN00140

C:Genetics:

A:Gene: At4g04620

C:Superfamily: hypothetical protein YBL078c

Query Match 35.0%; Score 219; DB 2; Length 122;
Best Local Similarity 36.8%; Pred. No. 2,1e-14;
Matches 42; Conservative 31; Mismatches 39; Indels 2; Gaps 2;

7 FKORRSFADCKEVOOIRDOHPSKIPVIERKGEKOLPVLDTKFLVPHVNSLVKI 66

Db 6 FKSINPLEMRMAESTRIKAKYPERVPIVER-AGOSIVPDIKKKKYLVPAIDLIGQFYV 64

QY 67 IRRRLQNPLOAFLLVNOHSMVSVSTPIADIYEOFKDEDEGLYMYVASQETFG 120

Db 65 VKRKIKGAEKAIIVFV-KNLTPTALMSATIEFHKDEDEGLYMTYSGENTFG 117

RESULT 8

T47875

Hypothetical protein T4C21.50 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 21-Jul-2000

C:Accession: T47875

R:Choiste, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattivico, L.; Artiguenave, F.; Sa

.M.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.

submitted to the Protein Sequence Database, March 2000

A:Reference number: Z24479

A:Accession: T47875

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-121 <CHO>

A:Cross-references: EMBL:AL162295

A:Experimental source: cultivar Columbia; BAC clone T4C21

C:Genetics:

A:Map position: 3

A:Initrons: 13/1; 31/3; 49/2; 89/1

A:Note: T4C21.50

C:Superfamily: hypothetical protein YBL078c

Query Match 34.7%; Score 217; DB 2; Length 121;
Best Local Similarity 37.0%; Pred. No. 3,3e-14;
Matches 44; Conservative 32; Mismatches 39; Indels 4; Gaps 3;

3 SDRPFKORRSFADCKEVOOIRDOHPSKIPVIERKGEKQLPVLDKTRFLVPHVNS 61

Db 2 SNVSFKODHDEPKKAKALIREKYSRVPIVE-KSEKSDIPNDKKITLVPAIDLIG 59

QY 62 ELVKIIRRLQNPLOAFLLVNOHSMVSVSTPIADIYEOFKDEDEGLYMYVASQETFG 120

Db 60 QFVYVIRKRIQLSSEKAFIFVVD-NVLPITGAMMSTIYDEKDEDEGLYMTYSGENTFG 117

RESULT 9

S45432
Hypothetical protein YBL078c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YBL0732

C:Species: Saccharomyces cerevisiae

C:Date: 09-Jun-1994 #sequence_revision 09-Sep-1994 #text_change 19-Apr-2002

C:Accession: S45432; S45814; S45819; S59229

R:Obermayer, R.; Gassenhuber, J.; Piravandi, E.; Domdey, H.

submitted to the EMBL Data Library, May 1994

A:Description: Sequence analysis of a 78.6 kb segment of the left end of Saccharomyces

A:Reference number: S45387

A:Accession: S45432

A:Molecule type: DNA

A:Residues: 1-117 <DB>

A:Cross-references: EMBL:X79489; NID:g496661; PIDN:CA456032.1; PID:g49795

A:Experimental source: Strain S288c

R:Contreras, R.; Piers, W.; Lough, M.; Molmans, F.

submitted to the Protein Sequence Database, August 1994

A:Reference number: S45802

A:Accession: S45814

A:Molecule type: DNA

A:Residues: 1-117 <CON>

A:Cross-references: EMBL:Z35839; NID:g536124; PIDN:CA484899.1; PID:g56125; MIPS:YBL0

A:Experimental source: Strain S288c

R:Domdey, H.; Gassenhuber, H.; Obermayer, R.; Piravandi, E.

submitted to the Protein Sequence Database, August 1994

A:Reference number: S45815

A:Accession: S45819

A:Molecule type: DNA

A:Residues: 1-117 <DB>

A:Cross-references: EMBL:Z35839; NID:g536124; PIDN:CA484899.1; PID:g56125; MIPS:YBL0

A:Experimental source: Strain S288c

R:Obermayer, R.; Gassenhuber, J.; Piravandi, E.; Domdey, H.

Yeast 11, 1103-1112, 1995

A:Title: Sequence analysis of a 78.6 kb segment of the left end of Saccharomyces cere

A:Reference number: S59184; MUID:96076635; PMID:7502586

A:Accession: S59229

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-117 <OHV>

A:Cross-references: EMBL:X79489; NID:g496661; PIDN:CA456032.1; PID:g49795

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1994

C:Genetics:

A:Gene: SCD:A017

A:Cross-references: SCD:S0000174

A:Map position: 2L

C:Superfamily: hypothetical protein YBL078c

Query Match 34.6%; Score 215; DB 2; Length 117;
Best Local Similarity 37.8%; Pred. No. 4e-14;
Matches 45; Conservative 32; Mismatches 32; Indels 10; Gaps 4;

3 SDRPFKORRSFADCKEVOOIRDOHPSKIPVIERKGEKQLPVLDKTRFLVPHVNS 61

Db 7 SEYFPEKKA-----ESERIALDFKKNRIPVIG-RAEKSDIPNDKKITLVPAIDLIG 58

QY 62 ELVKIIRRLQNPLOAFLLVNOHSMVSVSTPIADIYEOFKDEDEGLYMYVASQETFG 120

Db 59 QFVYVIRKRIQLSSEKAFIFVVD-NVLPITGAMMSTIYDEKDEDEGLYMTYSGENTFG 116

RESULT 10

JC7698

GEC1 protein - quinea pua

C:Species: Cavia porcellus (guinea pig)

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Nov-2001

C:Accession: JC7698

R:Venier-Magnin, S.; Muller, S.; Sallot, M.; Radom, J.; Musard, J.F.; Adami, P.; Duj

Biochem. Biophys. Res. Commun. 284, 118-125, 2001

A:Title: A novel early estrogen-regulated gene gec1 encodes a protein related to GABA

A:Reference number: JC7698; MUID:21266936; PMID:11374880

A:Accession: JC7698

A:Molecule type: mRNA

A:Residues: 1-117 <VEK>

THIS PAGE BLANK (U&PTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 09:50:23 ; Search time 24 Seconds

(without alignments)
237.093 Million cell updates/sec

Title: US-09-904-603-1

Perfect score: 625

Sequence: 1 MPSDRPFKORSRFAHCKEV.....EKDQCFLYWVYASQETFGF 121

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	555	88.8	125	MP3A_HUMAN
2	516	82.6	119	MP13_BOVIN
3	516	82.6	124	MP3B_HUMAN
4	516	82.6	124	MP13_MOUSE
5	516	82.6	141	MP13_RAT
6	341.5	61.0	130	YHM6_CAEBL
7	248	39.7	117	GEF2_HUMAN
8	222	35.5	184	STRP_LACB1
9	216	34.6	117	APG8_YEAST
10	194	31.0	123	YOD9_CAEBL
11	74.5	11.9	660	DNL10_CHLMU
12	73	11.7	260	DH10_ARATH
13	73	11.7	746	RNG3_SCHRO
14	72.5	11.6	224	YPOJG_BACSU
15	72.5	11.6	682	NPH1_CAEBL
16	72	11.5	140	APGB_HUMAN
17	72	11.5	141	APGB_MOUSE
18	72	11.5	483	Y045_MYCE
19	72	11.5	1046	POL_FENY1
20	71	11.4	392	KBL_BACSU
21	71	11.4	396	YG78_METJA
22	71	11.4	558	C5G_METJA
23	71	11.4	2048	RRPL_SENDE
24	70.5	11.3	278	RRPL_SENDE
25	70.5	11.3	1085	RRP2_METJA
26	70	11.2	718	RRP2_BOVIN
27	70	11.2	1272	SL53_BOVIN
28	70	11.2	2228	RRP2_MOUSE
29	70	11.2	2228	RRP2_MOUSE
30	70	11.2	2228	RRP2_SENDE
31	69.5	11.1	592	SYD_THETN
32	69	11.0	511	MAIK_WHEAT
33	69	11.0	704	T2D4_DROME

34	69	11.0	1311	FMK2_HUMAN	P51816	homo sapien
35	68.5	11.0	453	SYS_ARCEU	O28244	archaeobac
36	68	10.9	113	RS6_HRCAP	O8K918	homo sapien
37	68	10.9	587	CN3_HUMAN	O80117	homo sapien
38	68	10.9	1616	P201_MYGE	Q49429	mycoplasma
39	68	10.9	1741	RC01_GIALA	P25202	gialdia lam
40	67.5	10.8	257	SPHR_SYNP7	P39663	synochloroc
41	67.5	10.8	288	Y489_RICPR	Q92457	tricketstia
42	67.5	10.8	1189	POL_MAEVM	P10272	batodon endo
43	67	10.7	511	MAIK_BOVIN	Q90m23	homo sapien
44	67	10.7	511	MAIK_HORVD	P17156	homo sapien
45	67	10.7	718	SL53_CANPA	P31637	canis fam11

ALIGNMENTS

RESULT 1
MP3A_HUMAN STANDARD: PK: 125 AA.
AC 09H492:
DT 28-FEB-2003 (rel. 41, Created)
DT 28-FEB-2003 (rel. 41, Last sequence update)
DE Microtubule-associated proteins 1A/1B light chain 3A (MAP1A/MAP1B L1)
DE A) (MAP1A/1B light chain 3 A).
CN MAP1A3A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE-21638749; PubMed-11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.F., Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L., Bailey J., Barlow K.F., Bates K.N., Beare D.M., Beasley G.P., Bird G.P., Blakey S.E., Bridgman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Cole C.M., Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.K., Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garnier J., Graham D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall J.E., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson G.M., Johnson B., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Lehaeslahto M.H., Levesha M.A., Lloyd C., Lloyd D.M., Love J.D., Marsh V.L., Martin S.L., McMahon L.J., McIlroy K., McMurtry A.A., Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H., Rice C.M., Ross M.T., Scott C.E., Selva H.K., Shownkeen R., Sims S., Skuce C.D., Smith M.L., Soderlund C., Stewart C.A., Sulston J., Swann R.M., Symore N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Tomsen A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A., Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S., Rogers J.
RT *The DNA sequence and comparative analysis of human chromosome 20.
RL Nature 414:865-871(2001).
CC -!- FUNCTION: STRUCTURAL PROTEIN INVOLVED IN THE FILAMENTOUS CROSSED-BRIDGING BETWEEN MICROTUBULES AND OTHER SKELETAL ELEMENTS (by similarity).
CC -!- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, L1, L2 AND L3, CAN ASSOCIATE WITH MAP1A AND MAP1B PROTEINS (by similarity).
CC -!- SIMILARITY: BELONGS TO THE MAP1 L3 / GEF2 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation in the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: AL118520; CAC14078.1; -
DR Genew: HGNC:6838; MAP1LC3A.
DR MIM: 601242; -
DR InterPro: IPR004241; MAP1_LC3.
DR Pfam: PF02991; MAP1_LC3; 1.
SQ SEQUENCE 125 AA: 14493 MW: 097697B5424FC425 CRC64:

Query Match 88.8%; Score 555; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 4.1e-49;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 ADRCKEVOQIRDOHPSKIPVIERKGEKQLPVLDTKFLVDPDHVMSLKYIRRLQL 73
DB 18 ADRCKEVOQIRDOHPSKIPVIERKGEKQLPVLDTKFLVDPDHVMSLKYIRRLQL 77
QY 74 NPTQAFLLVNOHSMVSVSTPIADIEQEKDEGFLYVYASQETFGF 121
DB 78 NPTQAFLLVNOHSMVSVSTPIADIEQEKDEGFLYVYASQETFGF 125

RESULT 2

MP13_BOVIN STANDARD: PRT; 119 AA.
ID MP13_BOVIN
AC 041515;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Microtubule-associated proteins 1A/1B light chain 3 (MAP1A/MAP1B LC3)
DE (MAP1A/1B light chain 3) (fragment).
GN MAP1LC3 OR MAP1ALC3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP TISSUE=Brain;
RC MEDLINE=94209331; PubMed=7908909;
RA Mann S.S., Hammarback J.A.;
RT "Molecular characterization of light chain 3. A microtubule binding
subunit of MAP1A and MAP1B."
RL J. Biol. Chem. 269:11492-11497(1994).
RN [2]
RP SEQUENCE OF 5-119 FROM N.A.
RX MEDLINE=98216781; PubMed=9557703;
RA Meyers G., Stoll D., Gunn M.;
RT "Insertion of a sequence encoding light chain 3 of microtubule-
associated proteins 1A and 1B in a pestivirus genome: connection with
virus cytopathogenicity and induction of lethal disease in cattle."
RL J. Virol. 72:4139-4148(1998).
CC -1- FUNCTION: STRUCTURAL PROTEIN INVOLVED IN THE FILAMENTOUS
CROSS-BRIDGING BETWEEN MICROTUBULES AND OTHER SKELETAL ELEMENTS.
CC -1- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE
WITH MAP1A AND MAP1B PROTEINS.
CC -1- SIMILARITY: BELONGS TO THE MAP1 LC3 / GEF2 FAMILY.
CC -1- CAUTION: REF.2 SEQUENCE ORIGINATES FROM STRAIN JACP OF PESTIVIRUS
TYPE 1 WHICH SEEMS TO CONTAIN A CELLULAR INSERTION OF PART OF
THE BOVINE HOST MAP1ALC3 GENE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: U08085; AAB72082.1; ALT_SEQ.
DR InterPro: IPR004241; MAP1_LC3.

DR Pfam: PF02991; MAP1_LC3; 1.
KW Microtubules.

FT CONFLICT 5 5 T -> P (IN REF. 1).
FT CONFLICT 11 11 T -> S (IN REF. 1).
FT CONFLICT 13 15 EDR -> DDV (IN REF. 1).
FT CONFLICT 17 20 EDVR -> KEVO (IN REF. 1).
FT NON_TER 119 119
SQ SEQUENCE 119 AA: 13998 MW: 72A21397AB26759 CRC64:

Query Match 82.6%; Score 516; DB 1; Length 119;
Best Local Similarity 83.2%; Pred. No. 3.3e-45;
Matches 99; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 2 PSDRPFKORSPFADRCKEVOQIRDOHPSKIPVIERKGEKQLPVLDTKFLVDPHVM 61
DB 1 PSDRPFKORSPFADRCKEVOQIRDOHPSKIPVIERKGEKQLPVLDTKFLVDPHVM 60
QY 62 ELVKIIRRLQLNPTQAFLLVNOHSMVSVSTPIADIEQEKDEGFLYVYASQETFG 120
DB 61 ELVKIIRRLQLNPTQAFLLVNOHSMVSVSTPIADIEQEKDEGFLYVYASQETFG 119

RESULT 3

MP3B_HUMAN STANDARD: PRT; 124 AA.
ID MP3B_HUMAN
AC 096208;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Microtubule-associated proteins 1A/1B light chain 3B (MAP1A/MAP1B LC3
B) (MAP1A/1B light chain 3 B).
GN MAP1LC3B OR MAP1ALC3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung carcinoma;
RA Pietas A., Petersen I., Schlens K., Petersen S.;
RT "Human cDNA of MAP1A/1B LC3 gene which is downregulated in lung
carcinoma cells."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Pheochromocytoma;
RA Peng Y., Li Y., Gu W., Xu S., Han Z., Fu G., Chen Z.;
RT "A novel gene expressed in human pheochromocytoma."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

CC SEQUENCE FROM N.A.
RA Kawabata A., Hiki J., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Kikuchi R., Ota T., Suzuki Y., Ogasashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "MEDO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: STRUCTURAL PROTEIN INVOLVED IN THE FILAMENTOUS CROSS-
BRIDGING BETWEEN MICROTUBULES AND OTHER SKELETAL ELEMENTS (BY
similarity).
CC -1- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE
WITH MAP1A AND MAP1B PROTEINS (BY similarity).
CC -1- SIMILARITY: BELONGS TO THE MAP1 LC3 / GEF2 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: AF303888; AAG23182.1; -
DR EMBL: AF183417; AAG09686.1; -

```

DR EMBL: AK025556; BAB15169.1;
DR Genew: HGNC:1352; MAP1LC3B.
DR InterPro: IPR004241; MAP1LC3.
DR Pfam: PF02991; MAP1LC3; 1.
KM Microtubules.
FT INITMET 0 0 BY SIMILARITY.
SQ SEQUENCE 124 AA: 14557 MW: BB147B1C5DA353F CRC64:
Query Match 82.6%; Score 516; DB 1; Length 124;
Best Local Similarity 82.4%; Pred. No. 3,4e-45;
Matches 98; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

OY 2 PSDFPFKQKSFADCKEYVQIRDPHSKIPVILIERKKGKQLPVLDTKFLVPHVYNS 61
   1 PSEKIFKQKSFEGKRVHVLREQHTKIPVILIERKKGKQLPVLDTKFLVPHVYNS 60
DB 1 ELKIIIRRLQNLNQAFFLLVNGHSMVSVSTPISEYSESEKEDGFLVWYASQETFG 119
   61 ELKIIIRRLQNLNQAFFLLVNGHSMVSVSTPISEYSESEKEDGFLVWYASQETFG 119

Db 61 ELKIIIRRLQNLNQAFFLLVNGHSMVSVSTPISEYSESEKEDGFLVWYASQETFG 119

RESULT 4
MAP1_MOUSE
ID MAP1_MOUSE STANDARD: PRT: 124 AA.
AC Q6GVV6; Q6DIR0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Microtubule-associated proteins 1A/1B light chain 3 (MAP1A/MAP1B LC3)
DE (MAP1A/1B light chain 3).
GN MAP1LC3 OR MAP1ALC3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo, Heart, and Kidney;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamamada I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli P., Barsh G.,
RA Blake J., Boffelli D., Bolunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RI Functional annotation of a full-length mouse cDNA collection.;
RI Nature 409:685-690(2001).
CC -1- FUNCTION: STRUCTURAL PROTEIN INVOLVED IN THE FILAMENTOUS CROSS-
CC BRIDGING BETWEEN MICROTUBULES AND OTHER SKELETAL ELEMENTS (BY
CC SIMILARITY).
CC -1- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE
CC WITH MAP1A AND MAP1B PROTEINS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE MAP1 LC3 / GEF2 FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcements/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AF255953; AAL83723.1;
DR EMBL: AK002795; BAB22364.1;
DR EMBL: AK003106; BAB22569.1;
DR EMBL: AK003205; BAB22641.1;
DR EMBL: AK003588; BAB22855.1;
DR EMBL: AK012604; BAB28350.1;
DR MGD: MGI:1914694; Map1lc3.
DR InterPro: IPR004241; MAP1LC3.
DR Pfam: PF02991; MAP1LC3; 1.
KM Microtubules.
FT INITMET 0 0 BY SIMILARITY.
SQ SEQUENCE 88 AA: 14486 MW: 520E4FF2AA93B78D CRC64:
Query Match 82.6%; Score 516; DB 1; Length 124;
Best Local Similarity 82.4%; Pred. No. 3,4e-45;
Matches 98; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

OY 2 PSDFPFKQKSFADCKEYVQIRDPHSKIPVILIERKKGKQLPVLDTKFLVPHVYNS 61
   1 PSEKIFKQKSFEGKRVHVLREQHTKIPVILIERKKGKQLPVLDTKFLVPHVYNS 60
DB 1 ELKIIIRRLQNLNQAFFLLVNGHSMVSVSTPISEYSESEKEDGFLVWYASQETFG 119
   61 ELKIIIRRLQNLNQAFFLLVNGHSMVSVSTPISEYSESEKEDGFLVWYASQETFG 119

Db 61 ELKIIIRRLQNLNQAFFLLVNGHSMVSVSTPISEYSESEKEDGFLVWYASQETFG 119

RESULT 5
MAP1_RAT
ID MAP1_RAT STANDARD: PRT: 141 AA.
AC Q6Z6Z5;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Microtubule-associated proteins 1A/1B light chain 3 (MAP1A/MAP1B LC3)
DE (MAP1A/1B light chain 3).
GN MAP1LC3 OR MAP1ALC3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=94209331; PubMed 7968909;
RA Mann S.S., Hammarback J.A.;
RT "Molecular characterization of light chain 3. A microtubule binding
RT subunit of MAP1A and MAP1B."
RL J. Biol. Chem. 269:11492-11497(1994).
CC -1- FUNCTION: STRUCTURAL PROTEIN INVOLVED IN THE FILAMENTOUS
CC CROSS-BRIDGING BETWEEN MICROTUBULES AND OTHER SKELETAL ELEMENTS.
CC -1- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE
CC WITH MAP1A AND MAP1B PROTEINS.
CC -1- TISSUE SPECIFICITY: ABUNDANT ONLY IN NEURONS. DETECTED IN TESTES.
CC -1- SIMILARITY: BELONGS TO THE MAP1 LC3 / GEF2 FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcements/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U05784; AAA20645.1;
DR PIR: A53624; A53624.
DR InterPro: IPR004241; MAP1LC3.
DR Pfam: PF02991; MAP1LC3; 1.

```

KW Microtubules. 0 BY SIMILARITY.
FT INIT_MBT 0
SQ SEQUENCE 141 AA: 16263 MW: F14E9D788D03D65D CRC64:
Query Match 82.68; Score 516; DB 1; Length 141;
Best Local Similarity 82.48; Pred. NO. 4e-45;
Matches 98; Conservative 12; Mismatches 9; Indels 0; Gaps 0;
QY 2 PSDRPFQRRSFARCKEVEQIIRDPQHSKIPVIERKYGKQVLDTKTKFLVDPHVM 61
DB 1 PSEKTFQRRSFARCKEVEQIIRDPQHSKIPVIERKYGKQVLDTKTKFLVDPHVM 60
QY 62 ELVKIIRRLQNLPTQAFLLVNGHSMVSTPIADIYDEKDEGFLYVYASQETFG 120
DB 61 ELIKIIRRLQNLPTQAFLLVNGHSMVSTPISEYVESRDEGFLYVYASQETFG 119
RESULT 6
YHM6_CAEEL STANDARD; PRT; 130 AA.
ID YHM6_CAEEL
AC Q23536;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 15.1 kDa protein ZK593.6 in chromosome IV.
GN ZK593.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peldoderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA McMuray A.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE MAP1 LC3 / GEF2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z69385; CAA93421.1;
DR PIR: T27920; T27920.
DR WormPep: ZK593.6; CE06627.
DR InterPro: IPR004241; MAP1_LC3.
DR Pfam: PF02991; MAP1_LC3; 1.
KW Hypothetical protein.
SQ SEQUENCE 130 AA: 15095 MW: 347518DF8C8D3154 CRC64:
Query Match 61.08; Score 381.5; DB 1; Length 130;
Best Local Similarity 58.38; Pred. NO. 1.3e-31;
Matches 70; Conservative 27; Mismatches 20; Indels 3; Gaps 1;
QY 1 MPDRPFQRRSFARCKEVEQIIRDPQHSKIPVIERKYGKQVLDTKTKFLVDPHVM 60
DB 14 VPS---FKRRPFQRRSFARCKEVEQIIRDPQHSKIPVIERKYGKQVLDTKTKFLVDPHVM 70
QY 61 SELVKIIRRLQNLPTQAFLLVNGHSMVSTPIADIYDEKDEGFLYVYASQETFG 120
DB 71 AELMSIVRRRLQNLPTQAFLLVNGHSMVSTPISEYVESRDEGFLYVYASQETFG 130
RESULT 7
GEF2_HUMAN STANDARD; PRT; 117 AA.
ID GEF2_HUMAN
AC O08765; Q9DCP8; Q9UQF7;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Ganglioside expression factor 2 (GEF-2) (General protein transport
DE factor p16) (GATE-16) (GABA(A) receptor-associated protein-like 2)
DE (MAP1 light chain 3 related protein).
GN GABARAPL2 OR GEF2 OR PLC3A.
OS Homo sapiens (Human).
OS Mus musculus (Mouse).
OS Rattus norvegicus (Rat), and
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606, 10090, 10116, 9913;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-Human;
RA Storch S., Braulke T.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES-Human;
RA Song H., Peng Y., Yu Y., Fu G., Mao M., Zhang Q., Zhu H., Li G.,
RA Luo M., Chen J., Hu R.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A., AND INTERACTION WITH UTK1.
RC SPECIES-Human; TISSUE-Frontal cortex;
RX MEDLINE=21066693; PubMed=11146101;
RA Okazaki N., Yan J., Yusa S., Ueno T., Komiant E., Masuo Y.,
RA Koga H., Muramatsu M.A.;
RL "Interaction of the Unc-51-like kinase and microtubule-associated
RL protein light chain 3 related proteins in the brain: possible role of
RL vesicular transport in axonal elongation.";
RN Brain Res. Mol. Brain Res. 85:1-12(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES-Human; TISSUE-Brain, and Testis;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES-Human, and Mouse;
RX MEDLINE=21309076; PubMed=11414770;
RA Xin Y., Yu L., Chen Z., Zheng L., Fu Q., Jiang J., Zhang P., Gong R.,
RA Zhao S.;
RL "Cloning, expression patterns, and chromosome localization of three
RL human and two mouse homologues of GABA(A) receptor-associated
RL protein.";
RN Genomics 74:408-413(2001).
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES-Rat; TISSUE-Brain;
RA Ogura K.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC SPECIES-Bovine; TISSUE-Brain;
RX MEDLINE=20211392; PubMed=10747018;
RA Savit Y., Legesse-Miller A., Porat A., Elazar Z.;
RL "GATE-16, a membrane transport modulator, interacts with NSF and the
RL Golgi v-SNARE GOS-28.";
RN EMBO J. 19:1494-1504(2000).
RN [8]
RP SEQUENCE FROM N.A.
RC SPECIES-Bovine; TISSUE-Kidney;
RX MEDLINE=22326861; PubMed=12438634;
RA Becher P., Thiel H.-J., Collins M., Brownlie J., Orlich M.;
RL "Cellular sequences in pestivirus genomes encoding gamma-aminobutyric
RL Acid (A) receptor-associated protein and Golgi-associated ATPase
RL enhancer of 16 kilodaltons.";
RN J. Virol. 76:13069-13076(2002).
RN [9]
RP SEQUENCE FROM N.A.
RC SPECIES-Mouse; STRAIN-C57BL/6J; TISSUE-Hippocampus, and Kidney;

Db 5 FKDEHPEKRRKAEARIRCKYPRIPVICK-ADRTDPTIDKKKLVPSDLTVGCFYV 63
QY 67 IRRRLQNLPTQAFLLVNOHSMVSVSTPIADIYEQEKDEDEGFLYVWYASQETFG 120
Db 64 IRRRIKLAPDKALFIFVDE-VLPPTALMSALYEHKEDNLFVYSGENTFG 116

RESULT 9
APG8_YEAST
ID STANDARD: PRT: 117 AA.
AC P38182;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Autophagy protein 8 [contains: Apg8FG].
GN APG8 OR AUT7 OR YBL078C OR YBL0732.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=96076635; PubMed=7502586;
RA Obermaier B., Gassenhuber J., Piravandi E., Domdey H.;
RT "Sequence analysis of a 78.6 kb segment of the left end of
RT Saccharomyces cerevisiae chromosome II.";
RL Yeast 11:1103-1112(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Contreras R., Fiers W., Logghe M., Molemans F.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=20493280; PubMed=11038174;
RA Kirisako T., Ichimura Y., Okada H., Kabeya Y., Mizushima N.,
RA Yoshimori T., Ohsumi M., Takao T., Ohsumi Y.;
RT "The reversible modification regulates the membrane-binding state of
RT Apg8/Aut7 essential for autophagy and the cytoplasm to vacuole
RT targeting pathway.";
RL Cell Biol. 151:263-276(2000).
CC -1- FUNCTION: Forms a protein complex with AUT7 to mediate attachment
CC of autophagosomes to microtubules.
CC -1- SUBCELLULAR LOCATION: Membrane-bound.
CC -1- PFM: The C-terminal Arg of APG8 is removed by APG4 to form APG8FG.
CC APG8FG forms a conjugate with an unidentified molecule "X" and
CC thereby binds tightly to membranes. The adduct APG8FG-X is
CC reversed to soluble or loosely membrane-bound APG8FG by cleavage
CC by APG4.
CC -1- SIMILARITY: BELONGS TO THE MAP1 LC3 / GEF2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X79489; CA56032.1; -;
DR EMBL: Z35839; CA48489.1; -;
DR PIR: S45432; S45432.
DR SGD: S0000174; AUT7.
DR GO: GO:0005875; C:Microtubule associated complex; IPI.
DR GO: GO:0008017; F:Microtubule binding activity; IPI.
DR GO: GO:0006914; P:autophagy; IGI.
DR GO: GO:0006623; P:Protein-vacuolar targeting; IMP.
DR Pfam: PF02991; MAP1_LC3; 1.
KM Autophagy; Membrane; Ubl conjugation pathway.
FT DOMAIN 1 116 APG8FG.
FT SITE 116 117 CLEAVAGE (BY APG4).
SQ SEQUENCE 117 AA; 13627 MW; 1246CEFF5A038819 CRC64;

Query Match 34.68; Score 216; DB 1; Length 117;
Best Local Similarity 37.88; Pred. No. 5.4e-15;
Matches 45; Conservative 32; Mismatches 32; Indels 10; Gaps 4;
QY 3 SDRPFKORRSFADRCKEVQOIRDOHPKIPVILIERYKGEK-OLPVLDTKTFVDPHVNMS 61
Db 7 SEVPEKRRKAEARIRCKYPRIPVICK-ADRTDPTIDKKKLVPSDLTVGCFYV 58
QY 62 ELKRIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYEQEKDEDEGFLYVWYASQETFG 120
Db 59 QFYVYVIRKRIKLAPDKALFIFVND-TLPPALMSALYEHKEDNLFVYSGENTFG 116

RESULT 10
YOD9_CAEEL
ID YOD9_CAEEL STANDARD: PRT: 123 AA.
AC Q09490;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C32D5.9 in chromosome II.
GN C32D5.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Miller N.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Kohara Y., Shin I T., Suzuki Y., Sugano S., Potdevin M.,
RA Thierry-Mieg Y., Thierry-Mieg D., Thierry-Mieg J.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE MAP1 LC3 / GEF2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U23511; AAC46797.1; -;
DR EMBL: AF326943; AAG49393.1; -;
DR PIR: T15740; T15740.
DR WormPep: C32D5.9; CE01849.
DR InterPro: IPR004241; MAP1_LC3.
DR Pfam: PF02991; MAP1_LC3; 1.
KM Hypothetical protein.
SQ SEQUENCE 123 AA; 14764 MW; BC08727A3101875A CRC64;

Query Match 31.08; Score 194; DB 1; Length 123;
Best Local Similarity 35.18; Pred. No. 9.4e-13;
Matches 40; Conservative 26; Mismatches 46; Indels 2; Gaps 2;
QY 7 FKORRSFADRCKEVQOIRDOHPKIPVILIERYKGEKOLPYLDTKTFVDPHVNMSLTKI 66
Db 5 YKFNENNEKRRKAEARIRCKYPRIPVICK-ADRTDPTIDKKKLVPSDLTVGCFYV 63
QY 67 IRRRLQNLPTQAFLLVNOHSMVSVSTPIADIYEQEKDEDEGFLYVWYASQETFG 120
Db 64 IRRRIKLAPDKALFIFVFN-NVDPQMTTMGOLYODHHEDEFLYIAYSDESVCYG 116

RESULT 11
DNLC_CHDMU
ID DNLC_CHDMU STANDARD: PRT: 660 AA.

AC 09PKP2:
 DI 16-OCT-2001 (Rel. 40, Created)
 DI 16-OCT-2001 (Rel. 40, Last sequence update)
 DI 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA Ligase (EC 6.5.1.2) (Polydeoxyribonucleotide synthase [NAD+]).
 GN LIGA OR TC0423.
 OS Chlamydia muridarum.
 NC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 NX NCBI_TaxID=83560;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mopn / N199;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heideberg J.F.,
 RA White O., Hickey E.R., Peterson J., Ulteback T., Berry K., Bass S.,
 RA Lither M., Weidman J., Khouri H., Craven R., Bowman C., Dodson R.,
 RA Ginn M., Nelson W., Deboy K., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser G.M.;
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 RT pneumoniae AR39.";
 RI Nucleic Acids Res. 28:1397-1406(2000).
 CC -1- FUNCTION: THIS PROTEIN CATALYZES THE FORMATION OF PHOSPHODIESTER
 CC LINKAGES BETWEEN 5'-PHOSPHORYL AND 3'-HYDROXYL GROUPS IN DOUBLE-
 CC STRANDED DNA USING NAD AS A COENZYME AND AS THE ENERGY SOURCE FOR
 CC THE REACTION. IT IS ESSENTIAL FOR DNA REPLICATION AND REPAIR OF
 CC DAMAGED DNA (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: NAD(+) + [deoxyribonucleotide(N) +
 CC (deoxyribonucleotide(N)-M) = AMP + nicotinamide nucleotide +
 CC (deoxyribonucleotide(N)-M).
 CC -1- SIMILARITY: BELONGS TO THE NAD-DEPENDENT DNA LIGASE FAMILY.
 CC -1- SIMILARITY: Contains 1 BRCT domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE002309; AAF39279.1; -;
 DR PIR: A81704; A81704.
 DR HSSP: 087703; 1B04.
 DR TIGR: TC0423; -;
 DR InterPro: IPR001357; BRCT.
 DR InterPro: IPR004150; DNA_Ligase_OB.
 DR InterPro: IPR001679; DNALigase.
 DR InterPro: IPR000445; HhH.
 DR InterPro: IPR003583; HhH_1.
 DR InterPro: IPR004149; Znf_DNALLigase_C4.
 DR Pfam: PF00533; BRCT; 1.
 DR Pfam: PF01653; DNA_Ligase_N; 1.
 DR Pfam: PF03120; DNA_Ligase_OB; 1.
 DR Pfam: PF03119; DNALigase_CBD; 1.
 DR Pfam: PF00633; HhH; 1.
 DR ProDom: PD003944; DNALLigase; 1.
 DR SMART: SM00292; BRCT; 1.
 DR SMART: SM00278; HhH; 1.
 DR SMART: SM00532; LIGANC; 1.
 DR TIGRFAMs: TIGR00575; dnlj; 1.
 DR PROSITE: PS50172; BRCT; 1.
 DR PROSITE: PS01055; DNA_LIGASE_N1; FALSE.NEG.
 DR PROSITE: PS01056; DNA_LIGASE_N2; 1.
 KW LIGase; DNA repair; DNA replication; NAD; Complete proteome.
 FT DOMAIN 581 660 BRCT.
 FT BINDING 114 114 AMP (BY SIMILARITY).
 SU SEQUENCE 660 AA; 73784 MW; 06BB89145E723EAD CRC64;
 Query Match 11.9%; Score 74.5; DB 1; Length 660;
 Best Local Similarity 24.5%; Pred. No. 7.5;
 Matches 26; Conservative 16; Mismatches 27; Indels 37; Gaps 3;
 4 DRPKRGRSFADKCKEVOIDRDQHS-KIPVIERYKGEKQPLVDKTKFLVPHVMNSE 62

DB 27 DRPTISDYSYDMKMEHLOKIEQHPEMKVSPITYLDCR-----
 QY 63 LKVIIRRLQNLPTVAFPLVNGHSMVSVPADIVTEQKDEDCF 108
 DB 67 -----PSGPIVWPSHFMVS-----IANVSLBLELEFF 96
 RESULT 12
 DH10_ARATH STANDARD; PRT; 260 AA.
 AC P42759;
 DI 01-NOV-1995 (Rel. 32, Created)
 DI 01-NOV-1995 (Rel. 32, Last sequence update)
 DI 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Dehydrin ERD10 (low-temperature-induced protein 11145).
 GN ERD10 OR LT145 OR LT129 OR AT1G20450 OR F5M15.21.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosales;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 NX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=94348723; PubMed=8069491;
 RA Kiyosue T., Yamaguchi-Shinozaki K., Shinozaki K.;
 RT "Characterization of two cDNAs (ERD10 and ERD13) corresponding to
 RT genes that respond rapidly to dehydration stress in Arabidopsis
 RT thaliana.";
 RI Plant Cell Physiol. 35:225-231(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Landsberg erecta; TISSUE=leaf;
 RX MEDLINE=96046757; PubMed=7579189;
 RA Wellin B.V., Olsson A., Palva E.T.;
 RT "Structure and organization of two closely related low-temperature-
 RT induced dhn/lea/rab-like genes in Arabidopsis thaliana L. Heynh.";
 RI Plant Mol. Biol. 29:391-395(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=11130712;
 RA Theologis A., Ecker J.R., Palm G.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman G.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Decker K.,
 RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hulzer L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam H.,
 RA Langan-Hooper S., Lee A., Lee J.M., Leuz G.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti K., Marziani A.,
 RA Miltner J., Miranda M., Nguyen M., Nierman W.C., Osborne B.L.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley I.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Ulteback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Xu G., Fraser G.M., Venter J.C., Davis K.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RI Nature 408:816-820(2000).
 RN [4]
 RP SEQUENCE OF 56-260 FROM N.A.
 RC STRAIN=cv. Landsberg erecta; TISSUE=leaf;
 RX MEDLINE=95035985; PubMed=7948863;
 RA Wellin B.V., Olsson A., Nylander M., Palva E.T.;
 RT "Characterization and differential expression of dhn/lea/rab-like
 RT genes during cold acclimation and drought stress in Arabidopsis
 RT thaliana.";
 RI Plant Mol. Biol. 26:131-144(1994).
 CC -1- TISSUE SPECIFICITY: IN STEMS, CAULINE LEAVES, ROOTS AND FLWERS;
 CC LOW LEVELS FOUND IN MATURING SEEDS. ABSENT IN DRY SEEDS.

CC -1- INDUCTION: BY DEHYDRATION, COLD STRESS AND ABSCESS ACID (ABA).
 CC INDUCTION AFTER 10 H. COLD STRESS INDUCTION PEAKS AT 1 H AND 5 H
 CC AFTER START OF COLD EXPOSURE.
 CC -1- SIMILARITY: BELONGS TO THE PLANT DEHYDRIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: D17714: BAA04568.1; -
 DR EMBL: X90958: CAA62448.1; -
 DR EMBL: AC027665: AAF79613.1; -
 DR EMBL: X77614: CAA54705.1; -
 DR PIR: S60480: S60480.
 DR InterPro: IPR000167: Dehydrin.
 DR Pfam: PF00257: dehydrin.1.
 DR PROSITE: PS00315: DEHYDRIN_1. 1.
 DR PROSITE: PS00823: DEHYDRIN_2. 2.
 KW Dehydrin; Repeat; Multigene family.
 FT DOMAIN 71 74 POLY-GLU.
 FT DOMAIN 88 92 POLY-GLU.
 FT DOMAIN 104 112 POLY-SER.
 FT DOMAIN 121 129 POLY-LYS.
 FT DOMAIN 183 186 POLY-LYS.
 FT DOMAIN 184 247 POLY-GLU.
 FT REPEAT 184 247 2 X 21 AA REPEATS, LYS-RICH.
 FT REPEAT 227 247 2.
 FT CONFLICT 56 56 H -> G (IN REF. 3).
 FT CONFLICT 144 144 Q -> E (IN REF. 3).
 SQ SEQUENCE 260 AA: 29548 MW: 31866FC6B4A1A5A CRC64;

Query Match 11.7%; Score 73; DB 1; Length 260;
 Best Local Similarity 24.8%; Pred. No. 3.5;
 Matches 28; Conservative 21; Mismatches 34; Indels 30; Gaps 5;

QY 8 KQRRSPADRCKEVQQRDQ-HRSKIYV-----IERKGEKQ--LPVLDKTFLYVD 56
 DB 141 EENQGVMDVIRKKEFPKPGDDVPPVVTMPAHSVEHDKPEEKEKGFMMKIKELPG 200
 QY 57 HVMSELYKIRRRRLQNLPTQAFELLVNQHSMVSVSTPIADIYDEKEDGFL 109
 DB 201 HSKKPEDSOV-----VNTTPIVETATPIADIPEEK---GFM 234

RESULT 13
 RN3_SCHPO STANDARD; PRT: 746 AA.
 AC 074994:
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ring assembly protein 3.
 GN RN3 OR SPC613.04C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.
 RC STRAIN=972;
 RX MEDLINE=20312839; PubMed=10852821;
 RT Wong K.C.Y., Nagai N.I., Iino Y., Yamamoto M., Balasubramanian M.K.:
 RT "Fission yeast Rng3p: an UCS-domain protein that mediates myosin II
 RT assembly during cytokinesis".
 RL J. Cell Sci. 113:2421-2432(2000).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 RA Rutherford K., Rutter S., Saunders R., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Welljens I., Vansteens E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambolt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Sapakowski G.V., Ussey D., Barrett B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe".
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: Essential for actinomyosin ring assembly during
 CC cytokinesis. Has a role in conjunction with F-actin, in
 CC assembling myosin II-containing proteins, such as myo2, at the
 CC division site.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; associated with the
 CC actinomyosin ring.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AJ011773: CAA09767.1; -
 DR EMBL: AL031644: CAA21056.1; -
 DR PIR: T43648: T43648.
 DR GeneDB_Spomb: SPC613.04C; -
 KW Cell division; Septation.
 SQ SEQUENCE 746 AA: 84471 MW: A0158CEDB110D71E CRC64;

Query Match 11.7%; Score 73; DB 1; Length 746;
 Best Local Similarity 23.9%; Pred. No. 12;
 Matches 28; Conservative 25; Mismatches 44; Indels 20; Gaps 6;

QY 7 FKQRRSPADR-CKEVOQIRDOHPKIPYIIE-----RKGEKQLPVLDKTFLYVDHVM 60
 DB 291 FKRCRAIIRKKESEP--IKDMHTRVAILISKNAFOYOLSOVPIPTL---KLLOSKEYD 345
 QY 61 SELVRIIRRRQLNLPTQAF-----LLVNQHSMVSVSTPIADIYDEKEDGFLVMY 113
 DB 346 SVLEALNLOSSTLGGVKQLIADDSCLLNLSKLDIOWISPL-----DASSIATPIY 397

RESULT 14
 YPJG_BACSU STANDARD; PRT: 224 AA.
 AC P42981:
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein ypjG.
 GN ypjG OR JOG.
 OS Bacillus subtilis.

```

000 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
001 NCBI_TaxID=1423;
002 [1]
003 SEQUENCE FROM N.A.
004 STRAIN=168 / Maidburg;
005 MEDLINE=96349105; PubMed=8760912;
006 Sotokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
007 Sotok P.;
008 "Sequence analysis of the Bacillus subtilis chromosome region between
009 the surA and kds loci cloned in a yeast artificial chromosome.";
010 Microbiology 142:2005-2016(1996).
011 [2]
012 SEQUENCE FROM N.A.
013 MEDLINE=96044033; PubMed=9384377;
014 STRAIN=168;
015 Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
016 Azevedo V., Berrero M.G., Bessieres P., Borotini A., Borcetti S.,
017 Borris R., Boursier L., Brans A., Braun M., Brigand S.C., Bron S.,
018 Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
019 Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,
020 Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
021 Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
022 Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
023 Gilm S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
024 Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
025 Hilbert H., Holstapel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
026 Joris B., Karamata D., Kasahara Y., Kleier-Blanchard M., Klein C.,
027 Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
028 Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
029 Lee S.M., Levine A., Liu H., Masuda S., Malet C., Medigue C.,
030 Medina N., Melillo R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
031 Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
032 Parro V., Pohl T.M., Portelle D., Portolillo S., Prescott A.M.,
033 Presenten E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
034 Rieger M., Rivoita C., Rocha E., Roche B., Rose M., Sadate Y.,
035 Sato T., Scanlan E., Schleich S., Schroeter R., Scottone F.,
036 Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
037 Sotokin A., Tacconi E., Takaki T., Takahashi H., Takemaru K.,
038 Takeuchi M., Tanakoshi A., Tanaka T., Tappara P., Tognoni A.,
039 Tosato V., Uchiyama S., Vandenberg M., Vannier F., Vassartoli A.,
040 Viari A., Wambuit R., Wedler E., Wedler H., Weitzenger T.,
041 Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
042 Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
043 "The complete genome sequence of the gram-positive bacterium Bacillus
044 subtilis.";
045 Nature 390:249-256(1997).
046 -1- SIMILARITY: TO F.COLI YAIS.
047 -----
048 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
049 between the Swiss Institute of Bioinformatics and the EMBL outstation -
050 the European Bioinformatics Institute. There are no restrictions on its
051 use by non-profit institutions as long as its content is in no way
052 modified and this statement is not removed. Usage by and for commercial
053 entities requires a license agreement (See http://www.isb-sib.ch/announce/
054 or send an email to license@isb-sib.ch).
055 -----
056 CC EMBL: 138424; AAA92876.1; -
057 EMBL: 147709; AAB38444.1; -
058 EMBL: 299115; CAB14163.1; -
059 EMBL: F69937; F69937.
060 Subtilisin; Bg11212; ypiG.
061 InterPro: IPR003737; DUF158.
062 Pfam: PF02585; DUF158. 1.
063 Hypothetical protein; Complete proteome.
064 KMW: 224 AA; 24806 MW; 95984EAF3F47090 CkC64;
065 SEQUENCE
066
067 Query Match 11.6%; Score 72.5; DB 1; Length 224;
068 Best Local Similarity 24.8%; Pred. No. 3.3;
069 Matches 29; Conservative 22; Mismatches 31; Indels 35; Gaps 6;
070
071 QY 24 RUDGHPK-----IVYIERKGEKQLPVLDKTK-----FLVPDHV-NMSEL 63
072 DB 107 KDRHPDGNAAALVEAIFSGAIKRYKDEKSLPAHKVSKYVYWMINGFHQPDFVIDISTD 166

```

```

073 QY 64 VKIIRRLQNPLOAFLLVNGHSMVSSTIPADLYEKGEDKEDFLVYVASELH 129
074 DB 167 IE-AKKOSLNATKSGH-----PSKIDSVSTPLI-----NGIIEVAREKVI 204
075
076 RESULT 15
077 NPH1_GAEEL
078 ID NPH1_GAEEL STANJARD; PRI: 682 AA.
079 AC 017972; Q21570;
080 DT 28-FEB-2003 (rel. 41, Created)
081 DT 28-FEB-2003 (rel. 41, Last sequence update)
082 DT 28-FEB-2003 (rel. 41, Last annotation update)
083 DE Nephrocystin-1 like protein.
084 GN M28.7
085 OS Caenorhabditis elegans.
086 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
087 Rhabditidae; Pelodierinae; Caenorhabditis.
088 OX NCBI_TaxID=6239;
089 [1]
090 SEQUENCE FROM N.A.
091 STRAIN=Bristol N2;
092 Gardner A.;
093 Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
094 -1- SIMILARITY: Contains 1 SH3 domain.
095 -1- SIMILARITY: BELONGS TO THE NEPHROCYSTIN 1 FAMILY.
096 -----
097 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
098 between the Swiss Institute of Bioinformatics and the EMBL outstation -
099 the European Bioinformatics Institute. There are no restrictions on its
100 use by non-profit institutions as long as its content is in no way
101 modified and this statement is not removed. Usage by and for commercial
102 entities requires a license agreement (See http://www.isb-sib.ch/announce/
103 or send an email to license@isb-sib.ch).
104 -----
105 CC EMBL: 249911; GAA90133.1; -
106 PIR: T23813; T23813.
107 Wormpep: M28.7; C616278.
108 InterPro: IPR001452; SH3.
109 Pfam: PF00018; SH3; 1.
110 SMART: SM00326; SH3; 1.
111 DR PROSITE: PS50004; SH3; 1.
112 DR Hypothetical protein; Coiled coil; SH3 domain.
113 FT DOMAIN 10 37 COILED COIL (POTENTIAL).
114 FT DOMAIN 43 100 COILED COIL (POTENTIAL).
115 FT DOMAIN 180 200 COILED COIL (POTENTIAL).
116 FT DOMAIN 215 275 SH3.
117 FT DOMAIN 155 173 GLN/PRO-RICH.
118 SEQUENCE 682 AA; 78795 MW; F30CCF8B4B552D CkC64;
119
120 Query Match 11.6%; Score 72.5; DB 1; Length 682;
121 Best Local Similarity 24.8%; Pred. No. 12;
122 Matches 25; Conservative 16; Mismatches 43; Indels 17; Gaps 4;
123
124 QY 10 RRFADFCKEYQVQIRDDHPKRI--PVIIKPKYKGLPVLDKTK-----FLVPDHV-NMSEL 64
125 DB 43 RHFVPCQCGHRLDDHRRNKEKAKIDELVKKKNALFQJDKLKRILALSPKEDLSFS 102
126 QY 64 VKIIRRLQNPLOAFLLVNGHSMVSSTIPADLYEKGEDKEDFLVYVASELH 106
127 DB 103 VSDVDSSEERK-----KMAALIGRKSTIMYNDSESD 134

```

Search completed: August 11, 2003, 10:14:45
Job time : 26 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 10:08:18 ; Search time 95 Seconds
(without alignments)
328.678 Million cell updates/sec

Title: US-09-904-603-1
Perfect score: 625
Sequence: 1 MPEQRPXKGRSFADRCKEY.....EKDEDGLYVYASQIFGCF 121

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_plant:*
- 10: sp_protist:*
- 11: sp_proteus:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	625	100.0	121	4 Q9BXW5	Q9BXW5 homo sapien
2	625	100.0	121	11 Q91VR7	Q91VR7 mus musculu
3	625	100.0	170	4 Q9H491	Q9H491 homo sapien
4	622	99.5	121	11 Q9DC74	Q9DC74 mus musculu
5	359	57.4	147	4 Q9BXW4	Q9BXW4 homo sapien
6	242	38.7	339	12 Q8B512	Q8B512 bovine vira
7	234	37.4	116	10 Q8H715	Q8H715 phytochtr
8	234	37.4	121	10 Q8VYK7	Q8VYK7 arabidopsi
9	230	36.8	120	10 Q8LBA9	Q8LBA9 arabidopsi
10	227	36.3	120	10 Q9SL04	Q9SL04 arabidopsi
11	227	36.3	122	10 Q945K6	Q945K6 arabidopsi
12	227	36.3	122	10 Q8LEW4	Q8LEW4 arabidopsi
13	226	36.2	122	10 Q8S928	Q8S928 arabidopsi
14	225	36.0	122	10 Q8S926	Q8S926 arabidopsi
15	224	35.8	121	3 Q94272	Q94272 schizosacch
16	223	35.7	115	10 Q9LRP7	Q9LRP7 arabidopsi

17	222	35.5	121	10 Q8L5P9	Q8L5P9 clect arad
18	221	35.4	121	3 Q8M7Y7	Q8M7Y7 neorosticta
19	221	35.4	121	3 Q8J262	Q8J262 podostemona
20	219	35.0	107	10 Q65447	Q65447 arabidopsi
21	219	35.0	122	10 Q9XEB5	Q9XEB5 arabidopsi
22	217	34.7	119	10 Q8S925	Q8S925 arabidopsi
23	217	34.7	119	10 Q8S927	Q8S927 arabidopsi
24	217	34.7	121	10 Q9L2Z9	Q9L2Z9 arabidopsi
25	208	34.3	111	6 Q8HYK6	Q8HYK6 bos taurus
26	208	34.3	129	4 Q9SQ05	Q9SQ05 arabidopsi
27	203	32.5	117	4 Q9H0R8	Q9H0R8 homo sapien
28	203	32.5	117	11 Q9J1Y7	Q9J1Y7 catia porca
29	203	32.5	239	12 Q8B513	Q8B513 bovine vira
30	201.5	32.2	149	10 Q8Q708	Q8Q708 arabidopsi
31	200	32.0	125	3 Q9N1J4	Q9N1J4 pichia past
32	199	31.8	117	11 Q8R4B8	Q8R4B8 mus musculu
33	194	31.0	117	4 Q9HYV6	Q9HYV6 homo sapien
34	186	29.8	121	5 Q9W2S2	Q9W2S2 diost pallia
35	184	29.4	124	5 Q8LJK2	Q8LJK2 plasmodium
36	183	29.3	117	4 Q951G6	Q951G6 homo sapien
37	183	29.3	117	6 Q8MKR8	Q8MKR8 oryctolagus
38	183	29.3	117	11 Q9Q017	Q9Q017 mus musculu
39	181	29.0	117	11 Q9J1T6	Q9J1T6 mus musculu
40	177	28.3	120	5 Q9VEG5	Q9VEG5 diost pallia
41	172	27.5	122	13 Q9DFN7	Q9DFN7 allitichthys
42	140	22.4	165	10 Q23496	Q23496 arabidopsi
43	92	14.7	829	17 Q8Q499	Q8Q499 pyrococcus
44	88.5	14.2	512	8 Q9GHG5	Q9GHG5 colicidia
45	85	13.6	841	17 Q59612	Q59612 pyrococcus

ALIGNMENTS

RESULT 1

Q9BXW5 PREDIMINARY: FR: 121 AA.

ID Q9BXW5: 01-JUN-2001 (TREMBLrel. 17, Created)

AC Q9BXW5: 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)

DE MAPI light chain 3-like protein 1 (Hypothetical protein).

GN DKFZP761L0515.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Yu L.;

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Uterus;

RA Strausberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Amygdala;

RA Ansoerg W., Winkler U., Mewes H.W., Weill B., Wiemann S.;

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF276658; AAK35151.1; -

DR EMBL: BC015810; AAH15810.1; -

DR EMBL: AL833855; CAD38714.1; -

DR InterPro: IPR004241; MAPI_1C3.

DR Pfam: PF02991; MAPI_1C3; 1.

KW Hypothetical protein.

SO SEQUENCE 121 AA: 14272 MW: 4801166957892AF3 CR64:

Query Match 100.0%; Score 625; DB 4; Length 121;
Best Local Similarity 100.0%; Prod. No. 176-56;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPDRPFKORRSFADCKEVOQIRDOHPSKIPVIERKYGKQPLVLDKTKFLVDPHVM 60
 DB 1 MSPDRPFKORRSFADCKEVOQIRDOHPSKIPVIERKYGKQPLVLDKTKFLVDPHVM 60
 QY 61 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYEQEKDEGFLVWYASQETFG 120
 DB 61 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYEQEKDEGFLVWYASQETFG 120
 QY 121 F 121
 DB 121 F 121

RESULT 2

Q91VR7 PRELIMINARY; PRT: 121 AA.
 AC Q91VR7;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2001 (TREMBLrel. 19, Last sequence update)
 DE MAPI light chain 3-like protein 1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBITaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast tumor;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC010596; ANH10596.1;
 DR InterPro: IPR004241; MAP1_LC3;
 DR Pfam: PF02991; MAP1_LC3; 1.
 SQ SEQUENCE 121 AA: 14272 MW: 48C1FBE8F7892AF3 CRC64;

Query Match 100.0%; Score 625; DB 11; Length 121;
 Best Local Similarity 100.0%; Pred. No. 1.5e-56;
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPDRPFKORRSFADCKEVOQIRDOHPSKIPVIERKYGKQPLVLDKTKFLVDPHVM 60
 DB 1 MSPDRPFKORRSFADCKEVOQIRDOHPSKIPVIERKYGKQPLVLDKTKFLVDPHVM 60
 QY 61 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYEQEKDEGFLVWYASQETFG 120
 DB 61 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYEQEKDEGFLVWYASQETFG 120
 QY 121 F 121
 DB 121 F 121

RESULT 3

Q9H491 PRELIMINARY; PRT: 170 AA.
 AC Q9H491;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DE BA346K17.1.2 (Novel protein similar to MAP1LC3 (Microtubule-associated
 DE proteins 1A/1B light chain 3) from Rat, Isoform 2) (Fragment).
 GN BA346K17.1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCBITaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Saira H.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL118520; CAC14079.1;
 DR InterPro: IPR004241; MAP1_LC3;
 DR Pfam: PF02991; MAP1_LC3; 1.

QY 1 MSPDRPFKORRSFADCKEVOQIRDOHPSKIPVIERKYGKQPLVLDKTKFLVDPHVM 60
 DB 1 MSPDRPFKORRSFADCKEVOQIRDOHPSKIPVIERKYGKQPLVLDKTKFLVDPHVM 60
 QY 61 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYEQEKDEGFLVWYASQETFG 120
 DB 61 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYEQEKDEGFLVWYASQETFG 120
 QY 121 F 121
 DB 121 F 121

FT NON TER 1 1
 SQ SEQUENCE 170 AA: 19189 MW: B8C7521EDC6E4823 CRC64;

Query Match 100.0%; Score 625; DB 4; Length 170;
 Best Local Similarity 100.0%; Pred. No. 2.3e-56;
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPDRPFKORRSFADCKEVOQIRDOHPSKIPVIERKYGKQPLVLDKTKFLVDPHVM 60
 DB 50 MSPDRPFKORRSFADCKEVOQIRDOHPSKIPVIERKYGKQPLVLDKTKFLVDPHVM 109
 QY 61 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYEQEKDEGFLVWYASQETFG 120
 DB 110 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYEQEKDEGFLVWYASQETFG 169
 QY 121 F 121
 DB 170 F 170

RESULT 4

Q9DC74 PRELIMINARY; PRT: 121 AA.
 AC Q9DC74;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE 1010001H21RIK protein.
 GN 1010001H21RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBITaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Heart;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikolaou I., Pesole G., Quackenbush J.,
 RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bona M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guslinich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,
 RA Lyons P., Marchionni L., Hashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata K., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
 RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK003122; BAB22582.1;
 DR MGI: MGI:1915661; 1010001H21RIK.
 DR InterPro: IPR004241; MAP1_LC3.
 DR Pfam: PF02991; MAP1_LC3; 1.
 SQ SEQUENCE 121 AA: 14286 MW: B8C1FBE8A2DF9A9E CRC64;

Query Match 99.5%; Score 622; DB 11; Length 121;
 Best Local Similarity 99.2%; Pred. No. 3.1e-56;
 Matches 120; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPDRPFKORRSFADCKEVOQIRDOHPSKIPVIERKYGKQPLVLDKTKFLVDPHVM 60
 DB 1 MSPDRPFKORRSFADCKEVOQIRDOHPSKIPVIERKYGKQPLVLDKTKFLVDPHVM 60
 QY 61 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYEQEKDEGFLVWYASQETFG 120
 DB 61 SELVKIIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYEQEKDEGFLVWYASQETFG 120
 QY 121 F 121
 DB 121 F 121

Db 65 VRRRIKLGAEKATFEV-KNTLPPTAALMSAIEEHNKDEGFLYMYSGENTFG 117

RESULT 14

ID 08S926 PRELIMINARY; PRT; 122 AA.

AC 08S926; 01-JUN-2002 (TReMBLrel. 21, Created)

DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)

DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)

DE Autophagy 8e.

GN ATAPG8E.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.

OX NCBI_TaxID=3702;

RN [1] SEQUENCE FROM N.A.

RA Hanaka H., Noda T., Shirano Y., Sato S., Kato T., Hayashi H.,

RT "Disruption of AtAPG9 indicates autophagy is involved in leaf

senescence and starvation tolerance in higher plant."

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB073179; BAB88391.1; -

DR InterPro: IPR004241; MAP1_LC3.

DR Pfam: PF02991; MAP1_LC3; 1.

SO SEQUENCE 122 AA; 13948 MW; 9C86933882DA3A1B CRC64;

Query Match 36.0%; Score 225; DB 10; Length 122;
Best Local Similarity 39.5%; Pred. No. 1.8e-15;
Matches 47; Conservative 29; Mismatches 31; Indels 12; Gaps 4;

QY 7 FKORRSFADRCVQOIRDOHPSKIPYIIERYKGEK-QLPVLDTKFLVDPHVMSELYK 65

Db 7 FKMDDFEKKKAGRIREKYPDRIPYIVE--KAKESEVPNIDKKRYLPSDLTVGOFFVY 64

QY 66 IIRRLQNLPTQAFLLVNOHSMVSVSTP---IADIYEOEKDEDEGFLVMVYASOENFG 120

Db 65 VIRKRIKLSAEKATFEV-----NVLPPTGELMSSVYEDKDEGFLYITYSGENTFG 118

RESULT 15

ID 094272 PRELIMINARY; PRT; 121 AA.

AC 094272; 01-MAY-1999 (TReMBLrel. 10, Created)

DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)

DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)

DE Putative autophagy protein.

GN SPBP8B7.24C.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

RN [1] SEQUENCE FROM N.A.

RP STRAIN-972h-;

RA Beck A., Reinhardt R., Lyne M., Rajandream M.A., Barrell B.G.;

RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL032684; CAA21809.1; -

DR GeneDB; SPBP8B7.24C; -

DR InterPro: IPR004241; MAP1_LC3.

DR Pfam: PF02991; MAP1_LC3; 1.

SO SEQUENCE 121 AA; 14155 MW; 964BC3EFE5FAE983 CRC64;

Query Match 35.8%; Score 224; DB 3; Length 121;
Best Local Similarity 37.7%; Pred. No. 2.2e-15;
Matches 43; Conservative 31; Mismatches 38; Indels 2; Gaps 2;

QY 7 FKORRSFADRCVQOIRDOHPSKIPYIIERYKGEKOLPVLDTKFLVDPHVMSELYK 66

Db 5 FKDDFSEKRRKTESQRIREKYPDRIPYICEKV-DKSDIALADKKKYLVPSSDLTVGOFFVY 63

QY 67 IIRRLQNLPTQAFLLVNOHSMVSVSTPIADIYEOEKDEDEGFLVMVYASOENFG 120

Db 64 IIRKRIKLSAEKATFEV-----NVLPPTAALMSTIYEHNKSEDEGFLYITYSGENTFG 116

Search completed: August 11, 2003, 10:15:28
Job time : 97 secs